

Radio Occultations: A Wave Theory Treatment

William G. Melbourne

**MONOGRAPH 6
DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES**

Radio Occultations: A Wave Theory Treatment

November 2003

The research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.



Chapter 1

Radio Occultation Using Earth Satellites Background and Overview

1.1 Introduction

This monograph is concerned with the phase and amplitude of an electromagnetic wave during a radio occultation episode. This wave has passed through an intervening medium from a distant emitter and arrives at a receiver. The receiver measures the phase and amplitude of the wave over the duration of the occultation episode. These measurement sequences can be used to infer physical properties about the intervening medium.

Radio occultation refers to a sounding technique in which a radio wave from an emitting spacecraft passes through an intervening planetary atmosphere before arriving at the receiver. The words “occultation,” “occulted,” or “occulting” imply that the geometry involving the emitter, the planet and its atmosphere, and the receiver changes with time. Although an occulting or eclipsing planet (or moon) usually is involved, the word has come in recent times to also include non-occulting events, for example, satellite-to-satellite sounding through the ionosphere or receiving a reflected wave from a reflecting surface. From the perspective of the receiver for the strictly occulting case, the emitter is seen to be either rising or setting with respect to the limb of the occulting planet. As the radio wave from the emitter passes through the intervening atmosphere, its velocity and direction of propagation are altered by the refracting medium. The phase and amplitude of the wave at the receiver consequently are altered relative to their values that would hold without the intervening medium or the occulting planet. As time evolves, profiles of the phase variation and the amplitude variation at the receiver are generated and

recorded by the receiver. These profiles provide information about the refractive properties of the intervening medium.

In seismology, an array of seismometers spread over some geographical extent is used to study the various types of seismic waves arriving at each seismometer from a remote earthquake. With the array, one can measure the differential arrival times at the different stations in the array and also the spectral properties of the various waves. The different paths followed by these various waves and certain physical properties of the medium through which they pass can be inferred from these observations. In a radio occultation, it is the kinematics of the emitter/receiver pair over the duration of an occultation episode that provides analogous information.

1.1.1 History of the Occultation Technique

An astrometric observation in the eighteenth century measuring the times of ingress and egress of a lunar occultation of a star was probably among the first scientific applications of the technique. The method of lunar distances, the common seaman's alternative to the relatively expensive chronometer in the eighteenth and nineteenth centuries for keeping accurate Greenwich time for longitude determination at sea, and even for calibrating marine chronometers over very long voyages, depended crucially on an accurate lunar ephemeris. This in turn depended on accurate astrometric observations of the moon relative to the background stars and a good dynamical theory for the lunar orbit. Much later, the limb of the moon has been used as a knife-edge to obtain the microwave diffraction pattern from certain quasars [1]. The fringe spacing and amplitude of this diffraction pattern provide information about the angular distribution of radiant intensity from these very compact radio sources. Also, planetary atmospheres have been studied by analyzing stellar refraction and scintillation effects that occur during the ingress and egress periods when the star is occulted by the planet [2–4].

The radio occultation technique to sound planetary atmospheres using spacecraft began almost at the dawn of the era of planetary exploration. The first spacecraft to Mars in 1964, Mariner 4, flew along a trajectory that passed behind Mars as viewed from Earth [5,6]. The extra carrier phase delay and amplitude variation observed on the radio link between Mariner 4 and the Earth-based radio telescopes as Mariner 4 passed behind Mars and emerged from the other side provided valuable density information about its very tenuous atmosphere and also about its ionosphere [7]. Since then a score of experiments involving planetary missions have been undertaken to study the atmospheres of almost all of the planets in the solar system, including several moons and the rings of Saturn [8–12].

1.1.2 Occultations from Earth Satellites

Sounding of the Earth's atmosphere and ionosphere using the occultation technique theoretically can be performed with any two cooperating satellites. Prior to the Global Positioning System (GPS) becoming operational, a few early radio occultation experiments from a satellite-to-satellite tracking link had been conducted. These included the link between the Mir station and a geostationary satellite [13] and between xx and xx [QA: pls fix] [14]. However, this monograph focuses primarily on the carrier phase and amplitude measured by a GPS receiver onboard a low Earth orbiting spacecraft (LEO) while tracking the navigation signals emitted by a GPS satellite during its occultation by the Earth's limb [15]. GPS/MET (Global Positioning System/Meteorology), an occultation experiment that flew on MicroLab-1 and launched in 1995, was the first occultation experiment using the GPS [16–19]. Although experimental, over 11,000 occultations were used from GPS/MET to recover refractivity, density, pressure, temperature, and water vapor profiles [19]. GPS/MET provided a definitive engineering proof-of-concept of the occultation technique, and its data set became an experimental platform for implementing improved tracking and data processing schemes on subsequent Earth satellites with GPS occultation capability. GPS/MET also provided a basis for assessing the scientific and societal value of the technique in such diverse applications as meteorology, boundary layer studies, numerical weather prediction (NWP), and global climate change. Since then the Challenging Minisatellite Payload (CHAMP) (2001) [20], Satellite de Aplicaciones Cientificas-C (SAC-C) (2001), and Gravity Recovery and Climate Experiment (GRACE) (2002) satellites have been launched, from which radio occultation observations are now more or less continually made [21]. These missions alone could return nearly 1000 occultations per day. Future operational missions are planned, such as the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) configuration of occultation-dedicated LEOs to be launched in 2005 [22–24]. This system will provide near-real-time sounding information from about 4000 globally distributed occultations per day, which will be assimilated into NWP programs. Also, other global navigation satellite systems (GNSSs), such as the Russian Global Navigation Satellite System (GLONASS), and future systems, such as the planned European system, Galileo, will broaden the opportunities for dedicated satellite-to-satellite occultation missions [25,26].

Figure 1-1 depicts in exaggerated form a typical occultation scenario involving a LEO and the GPS satellite constellation. For a setting occultation, the about-to-be occulted GPS satellite will be seen from the LEO to be setting with respect to the limb of the Earth. The duration of a typical analyzed occultation through the neutral atmosphere (from roughly 100-km altitude to sea level) is less than 100 s. A ray from this GPS satellite passes through the

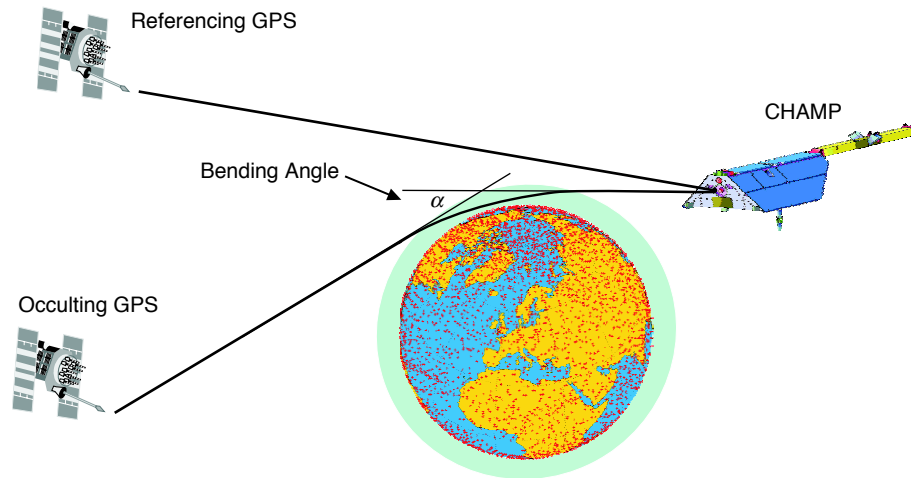


Fig. 1-1. Occultation scenario for a low Earth orbiter. From [91].

upper layers of the Earth's atmosphere at a near-horizontal rate of descent, skimming a deepest layer at its tangency point. A ray in this context may be defined by the normal to the cophasal surface of the carrier wave emitted by the GPS satellite. The ray then begins its near-horizontal rate of ascent through the upper layers, exits the atmosphere, and continues on to the LEO. The maximum total refractive bending of the ray is very small, about 1 deg from dry air. Additional refraction from water vapor, especially for tangency points in the lower troposphere, can double or triple that bending angle. A small bending-angle profile is a hallmark of a "thin" atmosphere. The index of refraction of dry air at microwave frequencies is about the same for light. At sunset the apparent sun for an observer on the ground is refracted through its own diameter, about 1/2 deg. The secular trend in the refractive gradient of dry air with altitude (which is near-exponential) is evident in the oblate shape of the apparent sun as its lower limb touches the horizon. Rays from the bottom limb usually are bent more than rays from the top limb. The fractured shape of the solar disk at some sunsets is caused by abrupt departures of the refractivity profile at low altitudes from the secular trend.

The ray arriving at the LEO from the occulted GPS satellite may not be unique, and indeed may not even exist in a geometric optics context. But for our purpose, we assume in Figure 1-1 that it does exist and that it is unique at the epoch of the observation. The inclination of the ray to the local geopotential surface is very slight; at 100 km from the tangency point it is about 1 deg in dry air, increasing linearly with ray path distance from the tangency point. At 400 km from the tangency point, the ray is about 10 km higher in altitude. Because of the near-exponential decrease in dry air density with height, it follows that most of the information about the atmosphere at a given epoch is

contained in a relatively narrow section of the ray about its tangency point, a few hundred kilometers in length [27].

As time evolves, the tangency point of the ray arriving at the LEO drifts deeper on average into the atmosphere for a setting occultation. Typically for the occultations selected for analysis, the initial cross-track angle of descent of the tangency point relative to the local vertical is less than 30 deg. The excess phase delay observed at the LEO, which is simply the extra phase induced by the refracting medium, will continue to increase on average over the course of the occultation because of the increasing air density with depth. This continues until the tangency point of the ray nears or contacts the planetary surface. The total excess delay can exceed 2 km with refractive bending angles up to 4 deg near the surface when water vapor is in abundance and the vertical gradient of its density is large. Defocusing and multipath, which tend to become strong in the lower troposphere, may reduce the signal amplitude to below a detection threshold before the ray contacts the surface, terminating the occultation episode several seconds or sometimes tens of seconds prematurely. However, sometimes the signal returns several seconds later, and sometimes in very smooth refractivity conditions even a knife-edge diffraction pattern from the limb or interference fringes from an ocean reflection are observed in the phase and amplitude before the direct signal from the GPS satellite is completely eclipsed [28,29]. On average, over the entire globe 80 percent of the CHAMP occultations reach to within 1 km of the surface, and 60 percent reach to within 1/2 km [21]. These encouraging statistics should improve in the future as new signal-tracking algorithms are implemented in the GPS receivers onboard the LEOs.

Figure 1-2 shows results from an early occultation from the GPS/MET experiment [18]. Here the excess phase delay of the L1 carrier in meters and its time derivative, excess Doppler in hertz, are shown over the last 90 seconds of the occultation. The lower abscissa shows coordinated Universal time (UTC), and its scale is linear. The upper abscissa shows the altitude of the ray path tangency point, and its scale is non-linear. The refractive gradient of the atmosphere increases markedly with depth, which effectively slows the average rate of descent of the tangency point because the refractive bending angle of the ray increases on average with depth. Therefore, the LEO must travel increasingly farther along its orbit to intercept these progressively deeper penetrating and more refracted rays. Near the Earth's surface, the average rate of descent of the tangency point typically is an order of magnitude smaller than it is in the upper atmosphere.

Figure 1-3 shows the amplitude of the L1 carrier for the same occultation [18]. The ordinate is the signal-to-noise ratio (SNR) in voltage, SNR_V , that would apply if the individual L1 carrier amplitude measurements were averaged over 1 s. The actual sample rate in this figure is 50 Hz. The averaging

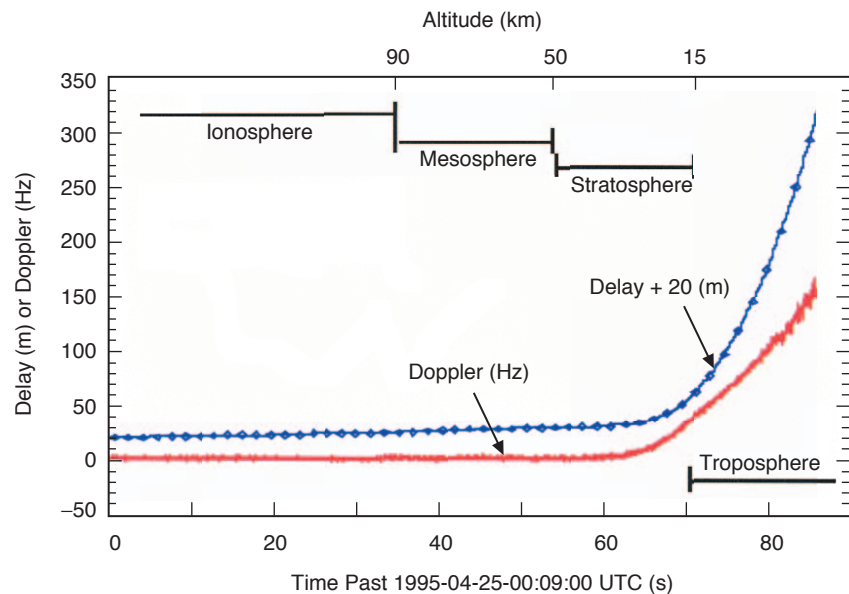


Fig. 1-2. Profile of excess Doppler and phase for a particular occultation of GPS PRN no. 28 observed by the GPS/MET experiment on MicroLab-1 on April 25, 1995, near Pago Pago. Redrawn from [18].

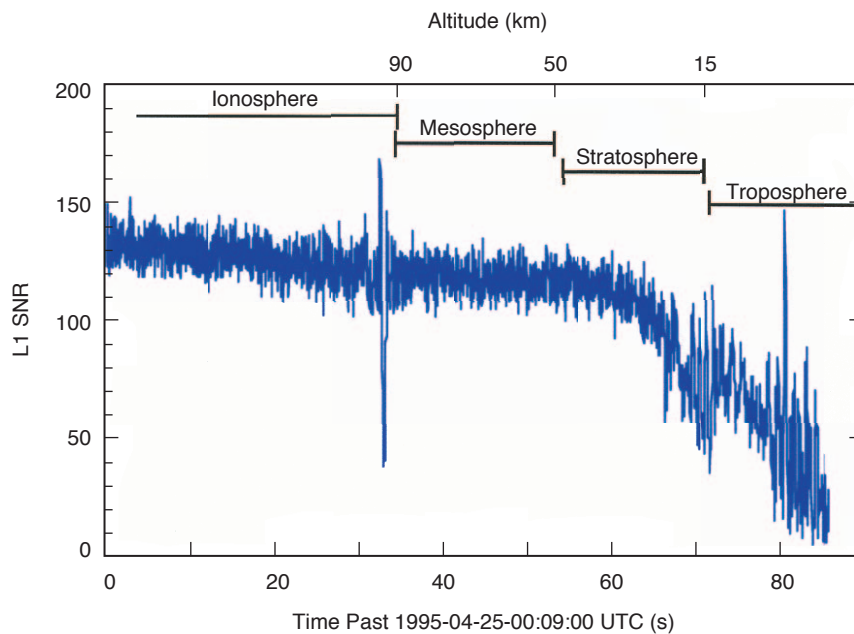


Fig. 1-3. Observed voltage SNR of the L1 carrier versus time for the same occultation shown in Fig. 1-2. Redrawn from [18].

time is 20 ms. Therefore, the thermal noise ($\sim 1/\text{SNR}_V$) for the points in this figure is $\sqrt{50}$ larger than the SNR_V values would imply. This figure clearly shows the secular defocusing caused by the increasing refractive gradient with depth. The refractive gradient disperses the directions of progressively deeper rays after exiting the atmosphere; the gradient effectively “de-collimates” the rays, thereby diluting their power at the LEO by spreading it over a larger area. The ratio of a small “collimated” area proportional to Δa , shown in Fig. 1-4(a), to the resulting de-collimated area proportional to $\Delta \sigma$ at the LEO is the defocusing factor. Appendix A derives a simple form for this factor:

$$\frac{\Delta a}{\Delta \sigma} = \zeta = \left(1 - D \frac{d\alpha}{da}\right)^{-1} \quad (1.1-1)$$

where D is effectively the distance of the LEO from the Earth’s limb and $d\alpha/da$ is the radial gradient of the refractive bending angle α . Figure 1-4(b) shows a relatively mild multipath scenario, including a shadow zone where $d\alpha/da$ is temporarily larger. Both the shadow zone and the interference from multipath waves cause the variability in SNR. The secular trend of SNR_V in Figure 1-3 is effectively given by $\zeta^{1/2}(\text{SNR}_V)_o$, where $\zeta^{1/2}$ is the defocusing from air and $(\text{SNR}_V)_o$ is the voltage SNR of the GPS signal that would be received at the LEO without the planet and its intervening atmosphere, the so-called “free-space” value. In addition to secular defocusing effects, this figure shows the strong transients in signal amplitude as the point of tangency of the ray passes through certain narrowly defined horizontal layers of the lower ionosphere and when it crosses the tropopause. These transients signal the presence of multipath-induced interference between different rays arriving concurrently at the LEO from different levels in the ionosphere and atmosphere. As the tangency point of the main ray cuts further down through successive layers of the atmosphere and into the middle and lower troposphere, the prevailing interference evident in this figure likely is induced by variable water vapor concentrations.

These transients observed in phase and amplitude raise a number of issues, some of which have been better dealt with than others. Significant progress on dealing with multipath has been made using back propagation and spectral techniques. Deep troughs in amplitude, shadow zones, and super-refractivity episodes, as well as caustics, are difficult for ray theory. How well these techniques work when the validity of ray theory itself is being strained is still to be established. The low SNR associated with many of these transients also makes it difficult to maintain connection in the LEO-observed phase measurements across them. We return to these topics later.

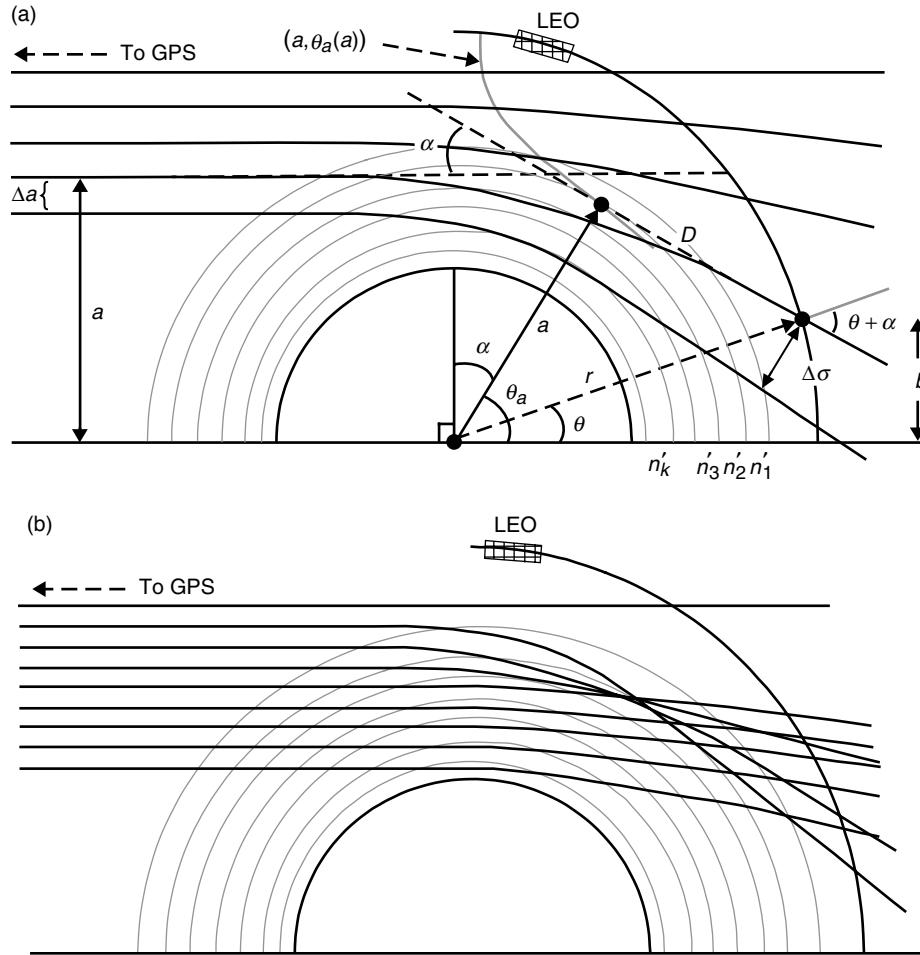


Fig. 1-4. Dispersive bending from the refractive gradient for an occultation observed from a LEO: (a) an ordered set of rays from an increasing refractive gradient with depth; no multipath, and (b) a non-monotonic refractive gradient results in a shadow zone and multipath.

1.1.3 The Global Positioning System

The GPS is operated by the U.S. Air Force (USAF). A system description of the GPS and the signal structure of its broadcast navigation signals are found in [30,31] and in many internet web sites. It suffices here to note a few details.

The GPS constellation is comprised of 24 satellites plus some on-orbit spares, more or less globally distributed up to 55-deg latitude. Their orbits are near-circular with a semi-major axis of about 4.1 Earth radii and with an inclination to the equator of 55 deg. A GPS satellite will be observed from a LEO to rise or set with respect to the Earth's limb on average once per 2 to 3

minutes, or at a rate of several hundred per day. The geographical distribution of the occultation tangency points is more or less global, but the actual distribution depends somewhat on the inclination of the LEO orbit plane and its altitude. For example, a LEO in a polar orbit returns fewer tangency points in the equatorial zone; fewer occulted GPS satellites in polar directions are viewed from that LEO because the GPS orbit inclination is only 55 deg. Figure 1-5 shows the global distribution of tangency points obtained over almost one month in 2001 for analyzed occultations from the CHAMP satellite. The lower density of points in the equatorial zone reveals CHAMP's near-polar orbital inclination. The geographical distribution also reveals vague clustering and striations resulting from commensurabilities between the satellite orbit periods. On roughly one-third of these occultations, a reflected ray from the Earth's surface also was detected [29].

The waveforms of the navigation signals broadcast by a GPS satellite are directional; they bathe the entire Earth with essentially full power. The 3-dB point of their radiant power distribution lobe is about 1400 km above the Earth's surface. Thus, a GPS receiver onboard a LEO, with an orbit radius typically well below this 3-dB threshold altitude, achieves about the same performance as a receiver on the ground.

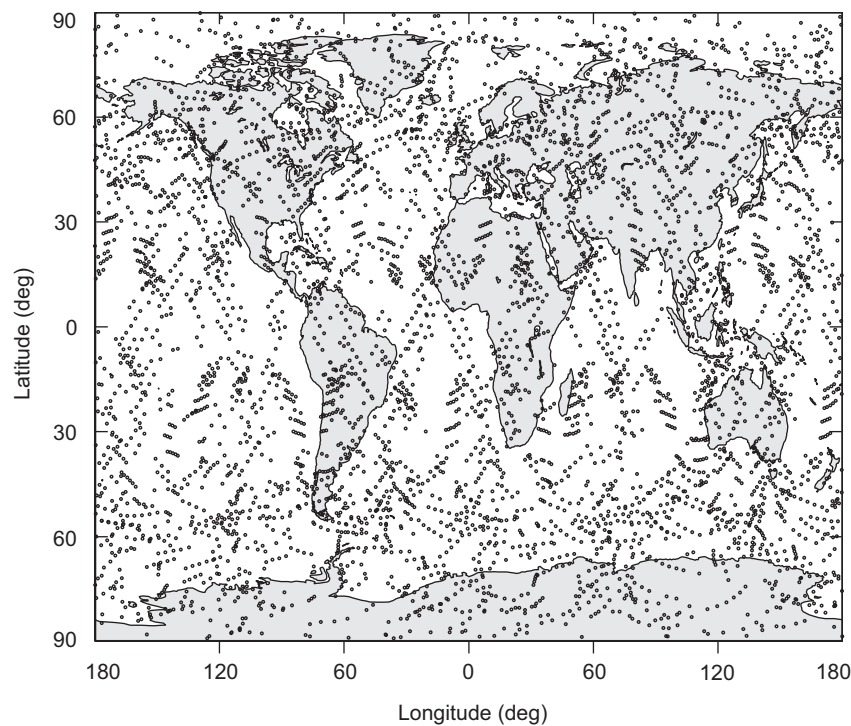


Fig. 1-5. Geographical distribution of nearly 4000 occultations obtained from CHAMP between May 14 and June 10, 2001. Redrawn from [29].

Each GPS satellite continuously broadcasts a set of square-wave codes by coherently modulating the phase of the carrier of the transmitted radio signal. These codes, which are unique to the broadcasting satellite, are pseudorandom and mutually orthogonal. They are used for ranging and for transmitting almanac and timing information. The mutual orthogonality property of the codes enables the receiver by cross-correlation techniques to isolate the received signals broadcast by a given satellite from all others, and to process in parallel the signals from all satellites in view of the receiver. The GPS satellites broadcast ranging codes on a pair of phase coherent L-band carriers, the L1 carrier at a frequency of 1575.4 MHz and L2 at 1227.6 MHz. These include an encrypted precision (P) code with a chip rate of 10.23 MHz on both carriers and the clear access or coarse acquisition (C/A) code at 1.023 MHz on the L1 carrier. The dual carriers are needed primarily to eliminate (or determine) the refraction effect from the ionosphere. For a microwave in the ionosphere, the refractivity is very nearly proportional to the local electron density and inversely proportional to the square of the carrier frequency. Therefore, the range and phase information received separately from the two carriers can be applied in concert to nearly completely decouple the ionospheric refraction effect by using this dispersive property of the ionospheric plasma. Newer versions of the GPS satellites planned for launch this decade will have an additional carrier at 1176.45 MHz (L5) and the C/A code also on L2. This will significantly improve receiver tracking operations using clear access ranging codes and increase the accuracy of the ionosphere calibration.

The signal structure of the ranging codes on a GPS signal is designed for near-real-time point positioning. By concurrently tracking four or more GPS satellites in diverse directions, the ranging code measurements can yield within a few seconds absolute point positions with an accuracy of roughly 10 m, and also one's time relative to the GPS clocks. One can obtain near-real-time relative positions (by concurrent tracking with two or more GPS receivers) with sub-meter or even sub-decimeter accuracy.

For occultation applications, however, one needs not the ranging information, but only very accurate measurements of the phase and amplitude of the L1 and L2 carriers, which are by-products of the range code tracking. For occultation applications, we may consider the radio signals arriving at a LEO from a GPS satellite as being a pair of spherical monochromatic waves from a distant point source at frequencies of 1575 MHz and 1228 MHz, respectively, plus Doppler shifts from kinematics and refraction of up to a few tens of kilohertz. Therefore, a high-performance GPS receiver used for space geodetic applications with millimeter-level accuracy requirements is a natural choice for occultation applications because it is designed to measure the phase of each carrier with sub-millimeter accuracy. Chapter 6 discusses certain additional aspects of such a space-rated receiver adapted for operations onboard a LEO.

1.1.4 Timing

A key factor for an accurate phase delay measurement is the epoch of the measurement. The measured phase at the receiver depends on the true phase accumulation between the emitter and the receiver, and the difference between the clock epochs of the emitter and receiver. Knowing the offset in time between the two clocks at a given instant is essential. More precisely, knowing the variability of this offset with time is essential. A constant offset is of no theoretical consequence (but it can be an operational problem) because the refraction information in the phase measurements is contained in their change with time. Each GPS satellite carries onboard up to four very precise cesium and/or rubidium frequency standards for controlling time and time intervals.

In 1999, the U.S. national policy was modified regarding certain operational aspects of the GPS. This change of policy led to the discontinuation of Selective Availability (SA) in May 2000, which had limited the accuracy of near-real-time point positioning to potential adversaries and to civilian users without access to decryption capability. By the late 1990s, SA became increasingly viewed as a cost and productivity issue for many GPS applications, military and civilian. As new technology and alternate means became available to the Department of Defense (DoD) for limiting access to the GPS, it became clear that, even though continuing SA provided a marginal defense benefit, it incurred an economic liability. SA deliberately degrades the near-real-time point-positioning accuracy of the GPS by at least an order of magnitude by causing the clock epoch errors in the GPS satellites to pseudo-randomly wander, nominally by the light-time equivalent of roughly 100 m over several minutes. Although the maximum deviation of the error is bounded and it can be averaged down substantially over 10 minutes or more, the short-term variability of SA poses a significant problem for clock epoch interpolation. SA dithers the onboard GPS master clock oscillator frequency at 10.23 MHz, which is derived from the atomic frequency standards and from which the chip rates of the codes and the frequencies of the L1 (154×10.23) and L2 (120×10.23) carriers are generated. The clock epochs depend on the integral over time of the oscillator frequency. The magnitude of the SA dithering has several possible levels of severity, which were set by certain alert or defense conditions. Over the years prior to 2000, SA had been set at a relatively low level but not at zero. Today SA is set to zero, but it has not been eliminated.

For carrier phase applications requiring high accuracy with SA turned on, multiple ground stations concurrently tracking all the GPS satellites in view at a relatively high sample rate of 1 Hz were needed [32,33]. This high-rate tracking was required to limit interpolation and/or extrapolation errors in rendering measured phases from different GPS satellites to a common epoch. Using these tracking data and applying a “double-differencing” or equivalent scheme among the tracked phase measurements referenced to common transmit epochs,

one can eliminate clock offset errors among the GPS satellites [34]. With these differencing operations, one incurs a penalty in increased effective thermal noise in the phase measurements, which becomes a concern in threshold situations where thermal noise becomes a limiting accuracy factor.

With SA off, the full inherent accuracy of the atomic frequency standards on the GPS satellites (with an effective 10-s frequency stability of a few parts in 10^{12}) can be realized with less costly ground tracking operations [35] and with more leisurely sample rates of 0.1 to 0.03 Hz. The interpolation error in the measured phase change over 10 s from GPS clock instability (with SA off) usually can be kept to well below 1 mm.

In addition to the GPS clock errors, the clock error in the LEO receiver also must be eliminated. However, the LEO usually carries an inferior frequency standard, which requires another strategy for eliminating this error source. This is depicted in Fig. 1-1, which shows the GPS receiver onboard the LEO observing the setting GPS satellite and simultaneously a second GPS satellite clear of any intervening medium, a so-called “clock” or reference satellite. (We assume that the dual-frequency phase measurements eliminate phase effects from the ionosphere.) Differencing the phase measurements from concurrent tracking of these two (now synchronized) GPS satellites at the same reception epoch eliminates the LEO clock error. The cost of this strategy in measurement precision is essentially a $\sqrt{2}$ increase in data noise, usually not a significant limitation except in threshold detection situations.

1.1.5 Ephemerides

For occultation applications, the knowledge error in the relative velocity between satellites must be controlled to a few tenths of a millimeter per second over the roughly 100-s duration of an occultation episode for the neutral atmosphere. This translates into a precision orbit determination (POD) requirement on the LEO of about 30 cm in accuracy, readily achievable these days with a GPS space geodetic receiver onboard. For the GPS constellation, the orbit information for each GPS satellite is included in its navigational signal and is maintained and operated by the USAF, usually with an accuracy of a few meters, or to a few hundred parts per billion. But for many scientific applications of the GPS, particularly space geodetic applications, one needs an accuracy of at most a few parts per billion. Pursuant to this goal, the International GPS Service (IGS) was inaugurated under the auspices of the International Union of Geodesy and Geophysics (IUGG) about a decade ago. The IGS consists of a globally distributed network of over 350 GPS ground tracking stations operated by almost as many different organizations, several communication and data information centers, data analysis centers, and a central bureau for oversight and user interfacing [36,37]. The IGS is comprised of over 200 organizations from over 80 countries collaborating on maintaining

and operating its various system elements. The scientific products from the IGS include very accurate and reliable ephemerides for the GPS satellites accurate to about 1 decimeter, plus a realization of a terrestrial reference frame for ground-based GPS receivers accurate to 1 to 2 parts per billion. These IGS-provided accuracies exceed the requirements for occultation applications.

1.2 Information Content in GPS Occultation Observations

A radio occultation observation profile consists of sequences of amplitude and phase measurements of the L1 and L2 carriers obtained by the LEO receiver over the course of an occultation episode. For the neutral atmosphere, a typical occultation contains a few thousand data points, depending on the sample rate. The measurement error on each point is statistically independent. Clock epoch errors in the transmitter and receiver have been removed using the differential tracking techniques described earlier. Using the POD information about the LEO and the GPS satellites, one can accurately calculate the phase accumulation from satellite kinematics. Subtracting this from the phase measurement leaves as a remainder the excess phase from refraction. Thus, the information content about the refracting medium is contained in the amplitude and excess phase measurement sequences, such as those shown in Figs. 1-2 and 1-3.

1.2.1 Connected Phase

An important aspect of the excess phase measurement sequence is its connectedness. To maintain an accurate phase profile over time, successive phase measurements must (or at least should) be connected, which means that the integer number of cycles accumulated between successive measurement epochs must be exactly accounted for in addition to the fractional cycle determinations made by the receiver at each epoch. In benign signal conditions, the receiver itself can produce connected phase measurements using an internal phase model based on previous phase measurements to extrapolate forward to the next measurement epoch. But in adverse conditions, the raw measurements must be supplemented with the application of more realistic models of the excess phase between epochs. Based on statistical studies from actual occultation profiles, the uncertainty in empirical excess Doppler models for the Earth's atmosphere is about 10 Hz, but with some outliers, which tend to become exacerbated deeper in the moist lower troposphere [34]. Thus, the maximum uncertainty in accumulated cycles between measurement epochs usually is about $10\Delta t$, where Δt is the separation time between successive epochs. For $\Delta t = 0.02$ s, the uncertainty in the integer cycle count usually is 0.2 or less. Fixing cycle breaks can be a significant off-line data-editing task, and sometimes flags are required to denote unfixable cycle breaks.

1.2.2 Sample Rate Versus Vertical Resolution

The canonical sample rate for the phase and amplitude measurements is 50 Hz, which is convenient for the receiver because that is the chip rate of the GPS header code. This is another phase-modulating, very low-rate, square-wave code on the L1 and L2 carriers that carries almanac, timing, health, and other information for the tracked satellite. The individual phase and amplitude samples can be averaged and reported at a coarser rate to reduce thermal noise effects, i.e., to increase the SNR, but at a cost of potentially poorer vertical resolution in the refracting medium. The average vertical velocity of a ray path tangency point is 2 to 3 km/s in the upper atmosphere and an order of magnitude smaller in the lower troposphere, where the refractive gradient is much stronger, and even slower where very large refractive gradients from water vapor occur. Thus, averaging the measurements over a time span broader than where thermal noise is no longer the limiting error source will further limit the vertical resolution with little benefit in measurement accuracy.

Resolution topics, such as the first Fresnel zone, and related topics are discussed further in Chapter 2 and Appendix A. The first Fresnel zone defines the resolution perpendicular to the ray path of a single occultation observation, much like the Airy disk provides a resolution threshold for a circular lens in an optical instrument. As with the Airy disk, the rays passing through this zone are more or less in phase and interfere “constructively”; rays outside this zone interfere destructively. In a refracting medium, the first Fresnel zone is elliptically shaped. The vertical radius is given by

$$\mathcal{F} = \sqrt{\lambda D |\zeta|} \quad (1.2-1)$$

where λ is the wavelength of the GPS carrier, about 20 cm; D is the limb distance of the LEO, roughly 3000 km; and ζ is the defocusing factor. The vertical width of the first Fresnel zone is on average about 1.5 km in the upper atmosphere and roughly 1/2 km in the lower troposphere from dry air alone. But near the Earth’s surface, it can be an order of magnitude smaller when strong refractive gradients from water vapor are present. It suffices here to note that the wave properties of the GPS signal, phase, and amplitude can be used with a sequence of observations to achieve a vertical resolution that is a small fraction of the width of the local first Fresnel zone, perhaps as small as 10 percent. A discussion of data-smoothing strategies involving the trade-off between resolution and measurement accuracy can be found in [34].

1.2.3 Inverting Radio Occultation Data

Although a wave-theoretic approach to recovering the refraction properties of the atmosphere is the principal theme of this monograph, it often is more

convenient and useful to cast propagation processes in terms of geometric optics. The following discussion is set in a ray-theoretic context.

The radio occultation profiles of phase and amplitude are analyzed, broadly speaking, in either a stand-alone mode or in an assimilation mode. In the stand-alone mode, the observation sequences are used to determine a profile for the index of refraction $n(\mathbf{r})$ from some sort of inversion process without preponderant reliance on a priori information. Usually some kind of symmetry or stratification geometry must be invoked for the medium to unambiguously determine $n(\mathbf{r})$. Effectively, the number of degrees of freedom in defining $n(\mathbf{r})$ must be no more than the number of independent observations in an occultation profile in order to achieve an unambiguous determination without using a priori information. For the Earth, local spherical symmetry that accounts for the Earth flattening usually is assumed. Local inhomogeneity in the refractivity along an equipotential surface usually is a small error source for the middle troposphere and higher, but water-vapor-induced horizontal gradients in refractivity near sea level can become a significant error source [38].

Figure 1-4 shows a schematic of one “onion layer” approach. Within each layer, the gradient of the index of refraction is assumed constant. Thus, we have a sequence of unknown parameters, n'_1, n'_2, \dots, n'_M , to be determined from M or more observations. At a specific epoch within the occultation episode, the ray traverses the atmosphere down to a minimum depth, the k th layer, but no deeper. Therefore, the observations for this ray will depend only on the k parameters n'_1, \dots, n'_k . If an unambiguous relationship holds between the observation at a given epoch and the ray—and this is a big if—then it is straightforward to form a linearized system of equations involving a triangular information matrix relating the observation sequence to these refraction parameters. The schematic in Fig. 1-4(a) depicts (with exaggerated bending) an ordered set of rays, i.e., no multipath, with a one-to-one relationship between the ray arriving at the LEO and the value of a , the impact parameter of the ray. With each successive measurement epoch in this figure, the ray at its tangency point passes through a new layer that is lower in altitude and that was not “sounded” before. In Fig. 1-4(b), we have multipath for some periods when more than one ray arrives at the LEO at the same time; the impact parameter of each ray can be moving either up or down with time, depending on the type of ray. Therefore, an analysis scheme must first be applied to the observations to discriminate among concurrently arriving rays.

In the linearized version relating observations to refraction parameters, an observation at a given epoch is the actual observation minus the predicted observation based on an initial estimate for parameter set $\hat{n}'_1, \hat{n}'_2, \dots, \hat{n}'_M$, and based on a provisional ray path passing through the medium that follows Snell’s law. Inversion of this information matrix, or inverting a weighted least-squares matrix version if over-determined, yields a determination of the

corrections $\Delta\hat{n}'_1, \Delta\hat{n}'_2, \dots, \Delta\hat{n}'_M$. One then updates $\hat{n}'_1, \hat{n}'_2, \dots, \hat{n}'_M$ by adding the corrections and generates a new ray path, forms new observation differences, and iterates until convergence is achieved. Assuming that $n_0 = 1, n'_0 = 0$, one can integrate this converged sequence $\hat{n}'_1, \hat{n}'_2, \dots, \hat{n}'_M$ downward to obtain a recovered profile $\hat{n}(r)$ for the index of refraction. As a practical matter, at very high altitudes, above 50 to 60 km, the noise in the occultation observable from measurement error or ionospheric calibration error becomes larger than the signal from the atmosphere. In these high altitudes, to initiate the downward integration process, one can statistically combine in a maximum-likelihood sense the noisy occultation observable with an a priori value furnished by a model atmosphere, along with the appropriate covariance matrices [39]. Because refractivity varies approximately exponentially with altitude at these heights, an error in this matching and handing-over process (from strong reliance on a priori information at higher altitudes to strong reliance on the data at lower altitudes) is damped out exponentially with decreasing altitude and is nearly gone within 2 scale heights.

When spherical symmetry applies and only a single ray arrives at the LEO over time, or when multiple rays can be discriminated according to their excess Doppler signature, then both the bending-angle and impact parameter values α and a are obtained from the derived excess Doppler information for each ray. Excess Doppler for a given ray is the time derivative of its excess phase. In this case, the Abel integral transform [see Eq. (1.2-5)] can be applied [40,41]. This integral transform directly recovers without iteration the profile $\hat{n}(r)$ from the observed bending-angle and impact parameter sequences for each evolving ray arriving at the LEO. At very high altitudes where the atmospheric signal is weak, the same kind of statistical hand-over using predicted excess Doppler from a model is still used to initiate the integral downward.

The accuracy of the recovery of the index of refraction can be determined theoretically based on an assessment of the error sources within the occultation system. It also can be inferred from statistical comparisons with NWP model values from, for example, the European Centre for Medium-Range Weather Forecasts (ECMWF) [19,42]. Fractional accuracy for the refractivity recovery of 0.1 to 1.0 percent is achievable from about 5 km up to 30 to 40 km. Here the assumption of local spherical symmetry usually is valid and the signal from the atmosphere is relatively strong compared to the measurement noise. But in the lower troposphere, horizontal variations in refractivity can become a significant error source when laminar symmetry is assumed, and tracking becomes more difficult and the measurements are noisier. These factors can limit further improvement with depth of the accuracy of the refractivity recovery.

If the chemical composition of the medium is known, one knows precisely the relationship between the density of the medium and its refractivity. In this case, one obtains the density profile $\rho(r)$ of the refracting medium from the

recovered profile for $n(r) - 1$. Invoking hydrostatic equilibrium for the medium in a gravity field $g(r)$, one has the hydrostatic equation for the pressure $dp/dr = -g\rho$. Integrating this yields the pressure profile $p(r)$. Invoking thermodynamic equilibrium allows one to apply the ideal gas law to the recovered density and pressure profiles to obtain the temperature profile.

For the Earth, the gaseous constituents of the neutral atmosphere affecting the refractivity at microwave frequencies are dry air and water vapor. For an air/water system in thermodynamic equilibrium, the Clausius–Clapeyron equation gives the partial pressure of water vapor, which varies exponentially with $-1/T$ [43]. Reducing the temperature by 10 percent reduces the saturated water vapor content by an order of magnitude. Below a threshold temperature of about 250 K, it is too cold for any residual water vapor to significantly contribute to the refractivity at L-band frequencies. The altitude corresponding to a temperature of 250 K typically is in the range 6 to 8 km. For altitudes above this threshold and below roughly 40 km, the potential accuracy of the temperature recovery is less than 1 K, probably less than 0.5 K for the 10- to 25-km range [21]. For higher atmospheric temperatures at lower altitudes, water vapor becomes an additional component in the refractivity. We speak of the “dry delay,” which is the phase delay caused by the refractivity from air, and the “wet delay,” which is the delay solely from water vapor. The latter arises from the permanent electric dipole moment of the water molecule, which is a significant contributor to the refractivity at microwave frequencies. The problems with the wet delay generally are that the density of water vapor is highly variable in space and time, its relative abundance or specific humidity is uncertain, and its refractivity per mole at L-band frequencies is much larger, about 16 to 18 times larger, than the refractivity per mole of dry air. Especially in the lower troposphere, this variability in specific humidity results in large refractive gradients that can double or even triple the refractive bending from dry air alone. In the stand-alone mode, therefore, the radio occultation technique requires ancillary information for those regions where the temperature is above 250 K to convert the recovered refractivity profile into unambiguous density, pressure, temperature, and specific humidity profiles. In temperate and tropical regions where the water vapor content is a major uncertainty, the temperature profile usually is relatively better known from NWP models, and it is often used to aid the occultation recovery of water vapor [19,21].

1.2.4 Assimilating Radio Occultation Data

In the assimilation mode, the occultation observable profiles are assimilated through statistical inference processes into a global atmospheric model where the basic meteorological quantities of the refracting medium, density, pressure, temperature, specific humidity, cloud, aerosol and precipitation distributions,

wind fields, and so on are defined. Therefore, the index of refraction $n(\mathbf{r})$ is already constrained to varying degrees in space and time by a usually much larger and more global data set in the atmospheric model. Covariance matrices characterizing the statistical aspects of the global data set and the atmospheric models controlling thermodynamic and dynamic processes are also part of this global system. In contrast to the stand-alone mode, no symmetry or stratification assumptions about the refracting medium are theoretically required here. Moreover, the granularity of the model for the refracting medium for numerical computation or, equivalently, the four-dimensional cell size (three in space and one in time) is already established and controlled by the continuity equation. Usually the number of degrees of freedom, even when constrained by the physical equations in the model, is vastly more than found in a typical stand-alone occultation model.

The 4-dimensional variational analysis (4DVAR) technique in NWP programs is a prime example of the assimilation mode for using the occultation data [44–47]. Here the occultation profiles are assimilated competitively in a statistical treatment of the data. The radio signal from the position of the occulted GPS satellite is propagated through this global model atmosphere with its free parameters provisionally fixed. At each observation epoch in the occultation sequence of M observations, the difference between an actual observable of the GPS signal (e.g., excess phase, excess Doppler or bending angle, signal amplitude) and its propagated value based on the provisional parameter set is computed. This difference becomes a component of an additional M -dimensional observation vector in the global data set. One adjusts the free parameters of the global model to minimize a cost function characterizing the goodness of fit of all the data weighted by their covariance matrix inverses, including the occultation data. The efficacy of the occultation information in this environment to modify the values of the free parameters obviously depends on the information content in the global system and on the assigned covariance matrices.

1.2.5 Rays and Stationarity

The Poynting vector, which defines the direction and magnitude of the radiant power in an electromagnetic wave at a given point in terms of the cross product of the electric and magnetic field vectors, is perpendicular to the cophasal surface of the wave at that point. The limiting form that the Poynting vector takes as the wavelength of the wave is shrunk to zero defines the tangent vector of the ray in geometric optics. The second-order differential equation that results from this limiting process gives the curvature of the ray at any point in terms of the gradient of the refractivity there. This limiting process also gives the amplitude of the ray based on the Poynting vector, its reflection and transmission properties that would apply to an electromagnetic wave of very

short wavelength, and so on. What is missing in ray theory, by definition, is diffraction. Also, ray theory (as geometric optics is usually defined) cannot deal adequately with caustic rays or with trapped rays in a super-refracting medium where the curvature of a ray, if it could exist in this environment, would temporarily be greater than the curvature of the local equipotential surface. A caustic ray is one that first comes in contact with the envelope to a family of rays, for example, a spectral component of a rainbow. Born and Wolf provide a comprehensive discussion of the foundations of geometric optics following this limiting approach in [48], which cites many references giving the historical development of geometric optics.

The same differential equation for the ray (the Euler equation) comes from the calculus of variations upon applying Fermat's principle (see Appendix A). The phase accumulation φ along a path through a three-dimensional refracting medium from point A to point B may be written in parametric form as

$$\varphi = k \int_A^B n(\mathbf{r}) ds = k \int_A^B n(\mathbf{r}) (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} dt \quad (1.2-2)$$

where $k = 2\pi / \lambda$ is the wave number of the wave, $n(\mathbf{r})$ is the index of refraction at the position $\mathbf{r} = \hat{x}x + \hat{y}y + \hat{z}z$, $ds = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} dt$ is the incremental arc length along the path, t is an arbitrary parameter denoting position along the path, $(\dot{}) = d() / dt$, and the triad $(\dot{x}, \dot{y}, \dot{z})$ defines the slope of the path. The path is described by the functions $x(t)$, $y(t)$, and $z(t)$, and we now stipulate that they describe a ray path. Fermat's principle requires that this path integral along a ray be stationary. In other words, the phase accumulation from the same initial point to the same end point along any other physical path neighboring the ray path would differ from the phase accumulation along the ray path in a second-order manner. Let the neighboring path be described by the functions $x(t) + \varepsilon \xi_x(t)$, $y(t) + \varepsilon \xi_y(t)$, and $z(t) + \varepsilon \xi_z(t)$, where ε is a small parameter. The functions $\xi_x(t)$, $\xi_y(t)$, and $\xi_z(t)$ are completely arbitrary other than that they must be physical and they must satisfy the boundary conditions, i.e., $\xi_x(t_A) = \xi_x(t_B) = 0$, etc. Then $d\varphi / d\varepsilon$ must be zero when evaluated along the ray, on which $\varepsilon = 0$. This stationary-phase condition requires that the ray path at every point satisfy the Euler differential equation or its integral equivalent where a discontinuity in $n(\mathbf{r})$ occurs. Usually the stationary value of φ along a ray is a local minimum, i.e., $d^2\varphi / d\varepsilon^2 > 0$, but not always. Anomalous rays provide a local maximum, and caustics involve a breakdown of the assumption of a non-zero second-order variation, which is embedded in the foundations of geometric optics.

The calculus of variations is a powerful technique for developing basic ray theory and many of its attributes. It may be used to develop the eikonal equation for obtaining phase delay along a ray, the transversality condition for

relating the change in end-point phase delay to small changes in end-point position (including relating excess Doppler to refractive bending angle), the Hamilton–Jacobi theory for obtaining the end-point phase delay expressed as a field variable in terms of the end-point position variables for a family of rays, and finally, for obtaining the canonical variable representation of a ray. We discuss these aspects in Chapter 2 and in Appendices A and B.

The stationary-phase property of rays affects one’s strategy for processing occultation data using ray theory, depending on one’s objective. In a sense, stationarity is both a blessing and a curse. Regarding the latter, consider the process of finding the ray path passing through the end points A and B for a given trial index of refraction profile $\hat{n}(\mathbf{r})$. One can forward propagate from A or backward propagate from B using Snell’s law and a ray-tracing method. By adjusting the slope parameters of the ray at these end points, one theoretically can converge on a complete ray path that matches the boundary conditions, if such a ray exists. The problem is that the calculated phase delay obtained by following this provisional path also must match the observed value φ for each observation epoch. This in turn requires adjustments in $\hat{n}(\mathbf{r})$ to reach the observed values for φ ; for the stand-alone mode, this would require an inversion process such as that described earlier. But along the ray, the phase delay is stationary with respect to small variations in the slope parameters or to variations in any other parameter, for example, the impact parameter, that characterizes an alternative nearby path that satisfies the boundary conditions. There is a singularity here. In effect, the “density” of admissible paths around the actual ray with a phase delay within the limits $(\varphi, \varphi + \delta\varphi)$, all satisfying the boundary conditions, becomes infinite along the actual ray where $\delta\varphi = 0$. This singularity in path density forms the basis of Fermat’s principle. The idea is that all physical paths in the vicinity of the actual ray have the same phase delay to first order. Therefore, the phasor, $\exp(i\varphi)$, for each path neighboring the ray is the same, to first order, and when the phasors from all of these possible nearby paths are summed (Huygen’s principle) and averaged to obtain the total field at point B , they sum constructively, reinforcing the amplitude. For paths with larger deviations from the ray, their first-order phase variation with respect to path deviation is not zero. Therefore, their phase delays for different paths in their neighborhood are distributed more or less randomly over 2π radians, and their phasors destructively combine. No average field results from those paths with a non-zero first-order variation, only fluctuations. This concept of constructive reinforcement for nearby paths about a ray and destructive reinforcement for paths away from the ray forms the basis for the idea of a ray, the basis for Fermat’s principle, and the basis for the first Fresnel zone. It also is the central concept in the Feynman sum-over-histories technique used in quantum electrodynamics to calculate the quantum wave function for the probability amplitude of a quantum event [49].

This singularity in path density results in numerical instability when the measured phase sequence per se is used as a criterion to reconstruct the most probable ray path for each observational epoch. On the actual ray, $d\varphi/d\varepsilon = 0$, where ε is a parameter describing the departure of a trial path from the actual ray path. By tinkering with ε , one tries to converge to the actual ray path. One would use an iterative differential correction scheme of the linear form $\delta\varepsilon = -(d\varphi/d\varepsilon)^{-1} \delta\hat{\varphi}$, where $\delta\hat{\varphi}$ is the observed value of φ minus the computed value $\hat{\varphi}$ based on the provisional profile $\hat{n}(\mathbf{r})$ and the provisional path through this medium between the end points. Here one attempts to determine iteratively the maximum-likelihood estimates of $\delta\varepsilon$, $\hat{n}(\mathbf{r})$, and the path. This scheme will run into computational trouble after successive iterations as the iterated path approaches the actual ray path because $d\varphi/d\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$.

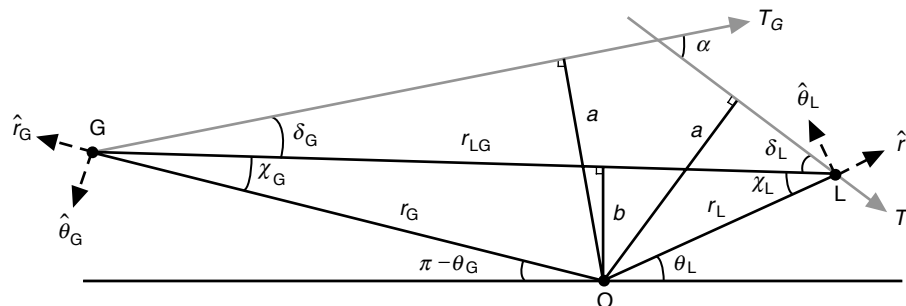
1.2.6 Excess Doppler

Stationarity also brings a blessing. If one slightly changes the end-point values of the ray, a new ray will pass through the new end points. The phase delay along the new ray between the new end points minus the phase delay along the old ray between the old points depends only on local conditions at the end points, not on the path in between. This is rather remarkable considering that the phase delay itself is computed from a path integral along a new path, which explicitly accounts for the variability in $\hat{n}(\mathbf{r})$ along the entire ray path. This local rather than global dependency on change is a unique property of a stationary-phase path.

Using the transversality condition, it is shown in Appendix A that the rate of change in excess phase resulting from the satellite velocities is given by

$$\lambda f_D = n_G \mathbf{T}_G \cdot \mathbf{V}_G - n_L \mathbf{T}_L \cdot \mathbf{V}_L - \dot{r}_{GL} \quad (1.2-3)$$

where $f_D = (d\varphi/dt)/2\pi$ is the rate of change of the excess phase in cycles, or excess Doppler shift of the ray, and λ is the wavelength of the harmonic wave. Here G denotes the position of the end point of the ray at the transmitting GPS satellite, and L denotes the end point position at the LEO. Also, \mathbf{V} is the velocity vector of the satellite at an end point of the ray, \mathbf{T} is the unit tangent vector of the ray at an end point, and \dot{r}_{GL} is the radial velocity between the end points G and L (see Fig. 1-6). POD information provides a determination of \mathbf{r}_G , \mathbf{V}_G , \mathbf{r}_L , and \mathbf{V}_L , and therefore \dot{r}_{GL} , at each observational epoch. It follows that the excess Doppler shift for a given ray, a quantity directly obtained from the observed excess phase profile for that ray, imposes a constraint on the unknown unit ray path tangent vectors \mathbf{T}_G and \mathbf{T}_L . If we assume coplanar propagation, usually a good assumption, then \mathbf{T} is defined by a single deflection angle



relative to the vector \mathbf{r}_{GL} between end points G and L (see Fig. 1-6). Ray tracing and Snell's law provide a second constraint on these two deflection angles. If we assume the satellites are out of the refracting medium, then $n_G = n_L = 1$. Thus, one determines the refractive bending angle, the angle α between T_G and T_L , provided that the ray is unique. No singularity arises in use of the excess Doppler to converge to the actual ray path from point G to point L.

Assuming that one has recovered the bending angle without ambiguity for a given ray, one obtains the impact parameter

$$a = n |\mathbf{r} \times \mathbf{T}| = nr \sin \gamma \quad (1.2-4)$$

for that ray, which is a constant along the ray when spherical symmetry applies. This is Bouguer's law, essentially Snell's law for a spherical symmetric medium. Here γ is the angle between the radius vector and the tangent vector of the ray. In ray theory, the constancy of this impact parameter a is the analogue of the conservation of angular momentum in classical mechanics. This is shown in Fig. 1-4. The impact parameter is related to the tangency point (r_*, θ_*) of the ray by the condition $a = r_* n(r_*)$. Then, as mentioned earlier, with the ordered series of values for the bending-angle and impact parameter pair (α_k, a_k) , $k=1, 2, \dots, M$, one can form the function $\alpha(\rho) = \alpha_k$, $\rho = a_k$, $k=1, 2, \dots, M$, and one can use the Abel transform

$$\log n(a) = \frac{1}{\pi} \int_a^\infty \frac{\alpha(\rho)}{\sqrt{\rho^2 - a^2}} d\rho \quad (1.2-5)$$

to recover the refractivity profile in terms of the impact parameter a . The Abel transform is derivable from the basic integral equation for the bending angle in a spherical symmetric medium (see Appendix A):

$$\alpha(a) = -2a \int_a^\infty \frac{d \log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - a^2}} \quad (1.2-6)$$

By substituting this form for $\alpha(a)$ into the integral operation on the right-hand side (RHS) of Eq. (1.2-5), one indeed recovers $\log n(a)$. If significant multipath occurs within a section of the observation series, then it should be noted that a_k , $k = 1, 2, \dots, M$ will not be time-ordered during that section. But, as long as one can successfully recover a sequence (α_k, a_k) that is ordered, then the Abel transform can be applied.

If spherical symmetry does not apply, then some ray tracing technique combined with the appropriate stratification model for the refractivity can be used to recover the refractivity profile, or one can use the assimilation mode.

Note from Eq. (1.2-6) that the existence of a ray between given end points with an impact parameter value of a requires that $\rho \geq a$ at all points along the ray. In a super-refracting layer, $d\rho / dr = (n + rn') < 0$, and this condition $\rho \geq a$ can be violated for a certain range of tangency points r_* . Within this range, no rays exist. Chapter 6 briefly discusses super-refractivity in a spherical shell.

1.3 Scientific Applications of GPS Occultation Observations

We briefly review the major uses of occultation observations, which fall broadly into the categories of meteorology, weather prediction, and global climate change. Accuracy, resolution, and the global distribution of the GPS occultations are key factors. Regarding accuracy, we have already noted the accuracy of the refractivity recovery from the stand-alone mode, 0.1 to 1 percent fractional accuracy above the lower troposphere, where the assumption of local spherical symmetry is likely to be satisfactory, and below 40 km. This translates into the same fractional accuracy range for the density determination for temperatures below 250 K. It also implies dekameter accuracy for a geopotential height determination of a point on a constant pressure surface, for example, the difference in the potential for the 300-mbar surface minus the potential of the ocean geoid, with the difference divided by the gravity coefficient [50]. It also translates for the same altitude range into sub-kelvin accuracy in temperature recovery.

Regarding these accuracy figures, the caveats about the applicability of spherical symmetry and the temperature being colder than 250 K have more to do with the model used and the analysis approach, and less to do with the inherent accuracy of the measurements. Good, clean phase and amplitude measurements down to the Earth's surface could yield 0.1-K accuracy for temperature recovery if we knew the water vapor content and the stratification. So, these caveats are somewhat misleading and tend to understate the inherent strength of the occultation data if our strategy is instead to assimilate them into a 4DVAR process.

Radio occultations bring new capabilities to the study of processes in meteorological models. With occultation data, one can study the fine structure of various refractive boundaries, such as the tropopause, a marine boundary, and a sporadic E-layer in the ionosphere. Temperature and pressure recovery from the middle troposphere through the middle stratosphere can lead to better understanding of energy transport and exchange processes across the tropopause and of how the atmosphere radiates dynamical energy and heat through acoustic gravity waves. Gravity waves are quite evident in occultation-derived temperature profiles in the stratosphere [51–54]. Phase and amplitude measurements made down to or near the Earth's surface provide powerful constraints on water vapor, which is such an important greenhouse gas and so relevant to energy transfer and balance processes in the atmosphere and to aerosol growth and cloud formation. Occultation-derived specific humidity distributions will provide sharper information about cyclogenesis processes, baroclinic wave and wind field development, and so on. More detail about many promising new capabilities is provided in [22,55].

1.3.1 Weather

The utility of occultation data when assimilated into 4DVAR programs depends greatly on the degree of complementarity of the information content in the occultation data set compared to the 4DVAR data set. The utility will be highest when other data are relatively sparse or when certain physical parameters of the 4DVAR model are less well constrained. This includes both data-sparse geographical locations and certain locations in altitude, even in data-rich geographical areas. Occultation profiles are in their best form for tangency points in the upper troposphere up to the lower stratosphere. Here the technique can provide unprecedented accuracy in density, pressure, and temperature, and it offers very sharp (perhaps dekameter-level) vertical resolution.

The utility of the occultation data also depends on their latency. The ephemeral aspects of the weather force very short latency requirements. That is why planned operational LEO occultation constellations, such as COSMIC,

have a 2- to 3-hour latency goal: occultation data to be completely assimilated within 3 hours of the actual occultation event.

Regarding geographical complementarity, consider again the planned COSMIC constellation, which will return on average about 10 occultations within a 3-hour period over a geographical area the size of the continental U.S. Obviously, 10 occultations from southern ocean or polar locations where weather data are sparse will have a bigger impact in weather models for those areas than will 10 occultations obtained over the continental U.S. and southern Canada in weather models for that area. In North America, the myriad weather sensors and observation programs is constraining.

Nonetheless, the water vapor information provided by the radio occultation data has a global utility. These data are powerful because they highly constrain the water vapor uncertainty. Accurate knowledge of specific humidity distributions, particularly in the lower troposphere, is key to more accurate weather prediction, for example, the ongoing development of wind fields, cloud formation, and precipitation in cyclogenesis processes. It is the water vapor information that makes reaching the Earth's surface with a high percentage of the radio occultations such an important goal for operational programs [56].

Pursuing this complementarity concept further, let us see how the water vapor information content in an occultation might fold into a 4DVAR process. The refractivity at L-band frequencies in the neutral atmosphere is given by [57,58]

$$N = n - 1 = \left(77.6 \frac{p}{T} + 3.73 \times 10^5 \frac{p_w}{T^2} \right) \times 10^{-6} + \text{negligible terms} \quad (1.3-1a)$$

where p is the pressure of the moist air in mbars, p_w is the partial pressure of the water vapor, and T is temperature in kelvins. Here and throughout this monograph, we use an unconventional definition for refractivity, $N = n - 1$, instead of the standard $N = (n - 1) \times 10^6$. Let us rewrite the refractivity equation in terms of the molecular number densities, n_a and n_w , in moles/m³, where n_a and n_w are the dry air and water vapor number densities, respectively. We use the ideal gas law $p_x = n_x RT$, where R is the universal gas constant. Then Eq. (1.3-1a) becomes

$$\left. \begin{aligned} N &= c(n_a + b(T)n_w) + \text{negligible terms} \\ c &= 6.45 \times 10^{-6} \text{ m}^3 \cdot \text{mole}^{-1}, \quad b(T) = 16.0 \left(\frac{300}{T} \right) + 1 \end{aligned} \right\} \quad (1.3-1b)$$

The relatively high sensitivity of the refractivity to the density of water vapor should be noted; the value of $b(T)$ typically is 16 to 18. Because the mean

molecular mass of water is 18 g/mole and dry air is 29 g/mole, the sensitivity of the specific refractivity is even greater.

Suppose now that we have inverted the occultation data to obtain an estimate $\hat{N}(r)$ of the refractivity. We could make a similar case for complementarity using a basic occultation observable, such as excess phase, Doppler, or amplitude. But it is easier to show using the recovered refractivity. Assume that at some epoch during the occultation episode the tangency point of the ray is located at a particular altitude where water vapor is a significant factor. Suppose that the error in the recovered refractivity there, $\delta\hat{N}$, is a random variable that is Gaussian distributed with a mean value of zero and a standard deviation of $\sigma_{\hat{N}}$. Then, from Eq. (1.3-1b), the probability density distribution $P[\delta\hat{N}]$ for $\delta\hat{N}$ written in terms of corresponding Gaussian errors in \hat{n}_a and \hat{n}_w is given by

$$P[\delta\hat{N}] = \frac{c^2 b}{2\pi\sigma_{\hat{N}}^2} \exp \left[-\frac{c^2 b^2}{2\sigma_{\hat{N}}^2} \left(\frac{\delta\hat{n}_a^2}{b^2} + \frac{\delta\hat{n}_w^2}{1^2} \right) \right] \quad (1.3-2)$$

Thus, a contour of constant probability density in $\delta\hat{N}$ is an ellipse with the semi-major axis along (or nearly so) the n_a axis and the semi-minor axis along the n_w axis. The coefficient $b(T)$ gives the ratio of these axes. The left-hand error ellipse, System A, in Figs. 1-7(a) and 1-7(b) gives qualitatively the contour for a constant probability density for $\delta\hat{N}$ from an occultation expressed in terms of the corresponding errors, $\delta\hat{n}_a$ and $\delta\hat{n}_w$, in the statistical estimates of n_a and n_w . For example, the ellipse might be the contour corresponding to the $1-\sigma$ values for $\delta\hat{n}_a$ and $\delta\hat{n}_w$; in this case, the probability of finding $\delta\hat{N}$ within the area inside of this ellipse is 40 percent.

Several comments can be made about Figs. 1-7(a) and 1-7(b). At this altitude of the tangency point and at this observation epoch, the error ellipse from the occultation information, System A, highly constrains the range of probable values for \hat{n}_w compared to the range of probable values for \hat{n}_a . System B in Fig. 1-7(a) shows another error ellipse for $\delta\hat{n}_a$ and $\delta\hat{n}_w$ from another information system. It is derived from a presumably higher-dimensional covariance matrix describing the uncertainties in the recovered values for the global parameter set in System B mapped to the same epoch and altitude and projected into our two-dimensional $(\delta\hat{n}_a, \delta\hat{n}_w)$ space. The information in System B is statistically independent from the information in System A. In Fig. 1-7(a), the information in System B is more or less “parallel” with the information in System A, nearly the same strength and the same weakness. In Fig. 1-7(b), System B is complementary to System A, more or less orthogonal, as characterized by the crossed error ellipses; the strength in

one system is a weakness in the other. The error ellipses for the two systems, A and B, may or may not be of comparable scale or eccentricity, but one can make an argument that they are often comparable. With the 2×2 covariance matrices for System A and System B, one can statistically combine the information in these two systems to obtain minimum variance estimates for n_a and n_w . The combined covariance matrix yields the System C minimum variance error ellipse. In Fig. 1-7(a), the combined System C, when the information in the two systems is comparable in accuracy, shows a roughly $1/\sqrt{2}$ improvement in probable errors all around, but it inherits the same strengths and weaknesses from the two systems. In the orthogonal case in Fig. 1-7(b), System C inherits the strengths from each system and none of the weaknesses; its error ellipse is dramatically improved. Water vapor content in NWP models is very difficult to know accurately in place and time, and it is generally less well-known in these models than are the total pressure and temperature. The near-orthogonal error ellipse in Fig. 1-7(b) for System B probably is more typical for NWP models.

In summary, in weather applications, complementarity from radio occultations, if they are timely, comes mainly from locations with relatively sparse meteorological data or from regions with more loosely constrained parameters. The other major contribution to NWP from the radio occultations is their highly constrained water vapor information, as just discussed. The long-perceived *bête noir* of the radio occultation technique for the Earth—water vapor—is in fact its strength.

1.3.2 Climate

In global climate change studies, detecting weak signatures in a noisy Earth system environment over decadal periods and longer is a major challenge. Here,

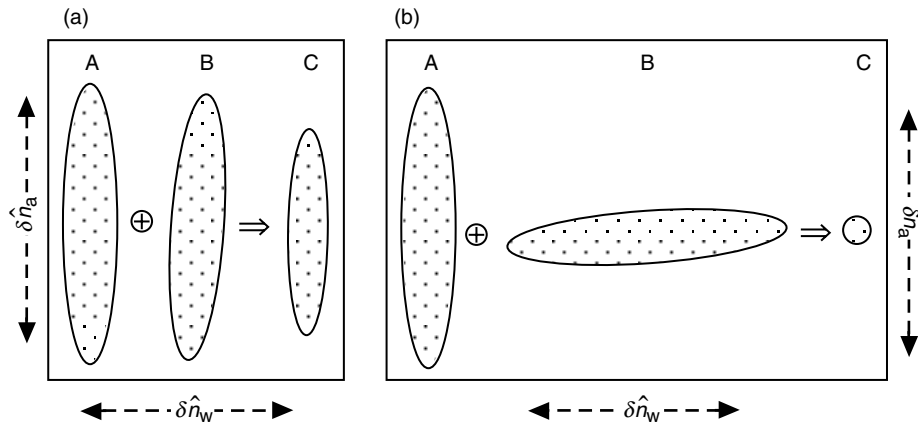


Fig. 1-7. Error ellipses for independent information systems A and B, and their minimum variance combination C: (a) parallel information and (b) complementary information.

the stability, high accuracy, and fine vertical resolution of the GPS occultations are key assets for use in climate studies. Except for the ionosphere, the GPS observations are essentially calibration-free, with no instrumental biases or calibrations that vary slowly. Even with the ionosphere, it would be a decadal variation in the usually small refraction effects from third- and higher-order frequency terms that could possibly “leak into” long-term observations in the stratosphere. This seems an unlikely scenario because modeling the ionosphere to aid the dual-frequency corrections significantly reduces errors from higher-order terms on the recovered refractivity in the stratosphere [59,60].

In monitoring possible secular trends, one can spatially average the recovered temperatures at a given altitude, or geopotential heights at a given pressure, to reduce by perhaps an order of magnitude the effect of random errors and variations. Thus, all the points lying within a specified geographical area, say a region 500×500 km in size, could be lumped and then averaged over time, weeks or months, enhancing the probability of detecting signatures with spectral power at very long periods.

Secular variations in the geopotential height of the 300-mbar surface, for example, can be a sensitive indicator of secular changes in the average temperature, \bar{T} , below. The sensitivity coefficient, $\partial h / \partial \bar{T}$, is about 30 m/K. A geopotential height change of 30 m at 300 mbar almost surely would be detectable over a long series of occultation observations, especially if it were regionally averaged, but ascertaining that it was a temperature signal rather than some other possible physical signal would remain a challenge.

A predicted secular effect associated with the global warming scenario is a warming troposphere but a cooling stratosphere. Therefore, monitoring averaged changes in recovered temperature profiles across the tropopause could be a very sensitive indicator of this warming trend. Other signatures associated with global warming to which averaged radio occultations might be applied are increased water vapor content and variability in the tropics, changes in the residual water vapor in the upper troposphere, and warming in the lower troposphere in polar regions. More examples and details on climate applications are given in [22,26].

1.4 Problems from Multipath and Some Remedies

The main problem from multipath, other than stressing the signal-tracking operations by the receiver, is in the ambiguity that it introduces regarding the appropriate excess Doppler value to use for each arriving ray. We have already noted in Fig. 1-3 the interference in amplitude from multipath rays for a particular occultation from GPS/MET. Figures 1-8(a) through 1-8(d) from CHAMP show similar interference in amplitude for a selected group of occultation profiles. The de-trended phase measurements for these occultations show similar interference. One should note from Fig. 1-8 the generally

improved SNR on CHAMP compared with GPS/MET. The GPS antenna on CHAMP has about 4 dB more gain, which on average yields about a factor of two increase in free space SNR_V . These profiles reveal striking and highly transient features in signal amplitude as the ray path tangency point descends through the atmosphere. Occultations passing through narrowly defined layers of the ionosphere also reveal similar transients in SNR and phase as a result of sharp gradients in electron density. Examples of passing through a sporadic E-layer in the lower ionosphere are seen in Fig. 1-8(b) and will be found later in Chapter 2 in Figs. 2-15 and 2-16 from GPS/MET.

A characteristic of many of these transients is their brevity, which, when the downward velocity of the ray is considered, translates into an altitude range that is usually narrow compared to the vertical width of the first Fresnel zone. Moreover, deep troughs in the SNR of the received signal combined with contiguous and abrupt flaring (and concomitant transients in phase) are often observed. These are almost sure signs of multipath and/or shadow zones, and even of ducting on occasion through a super-refracting layer [Fig. 1-8(a)]. In some instances, caustic rays are evident and also diffraction fringes when the physical refractivity feature is sharp enough and the neighboring environment is smooth enough.

Multipath is essentially the rule in occultations, not an exception. There are benign periods, of course, but even those usually show some small interference or scintillation that exceeds the thermal noise; it is a matter of degree. For example, the SNR in Fig. 1-8(c) at $t=0$ predicts a 1 percent thermal noise level ($\sqrt{50} / \text{SNR}_V$), but the actual noise is 2 to 3 times larger. One can average the samples down from a 50-Hz rate to a smaller rate, for example 1 Hz. Because the typical frequency difference, Δf , between interfering rays is several hertz, averaging dramatically improves the multipath “noise” compared to thermal noise. Over an averaging time $\Delta t \gg (\Delta f)^{-1}$, the amplitude of the fringes from multipath averages down as $1/\Delta t$, whereas thermal noise averages down as $(1/\Delta t)^{1/2}$. On the other hand, averaging can degrade the vertical resolution potential. For a LEO, the vertical motion of the ray path tangency point spans the vertical diameter of the local first Fresnel zone in a time interval of $\Delta T \approx 0.6\zeta^{-1/2}$ s, where ζ is the defocusing factor, unity at the top of the atmosphere and usually 0.1 to 0.01 at sea level. Therefore, averaging observations over a time interval longer than ΔT exceeds the Fresnel limit, which is the resolution from only a single observation. Sub-Fresnel vertical resolution, obtained from an ensemble of observations by means analogous to the resolution obtained from synthetic aperture radar systems, and a holy grail for boundary studies, can be considerably hampered by time-averaging the observations. Ultimately good SNR is the key.

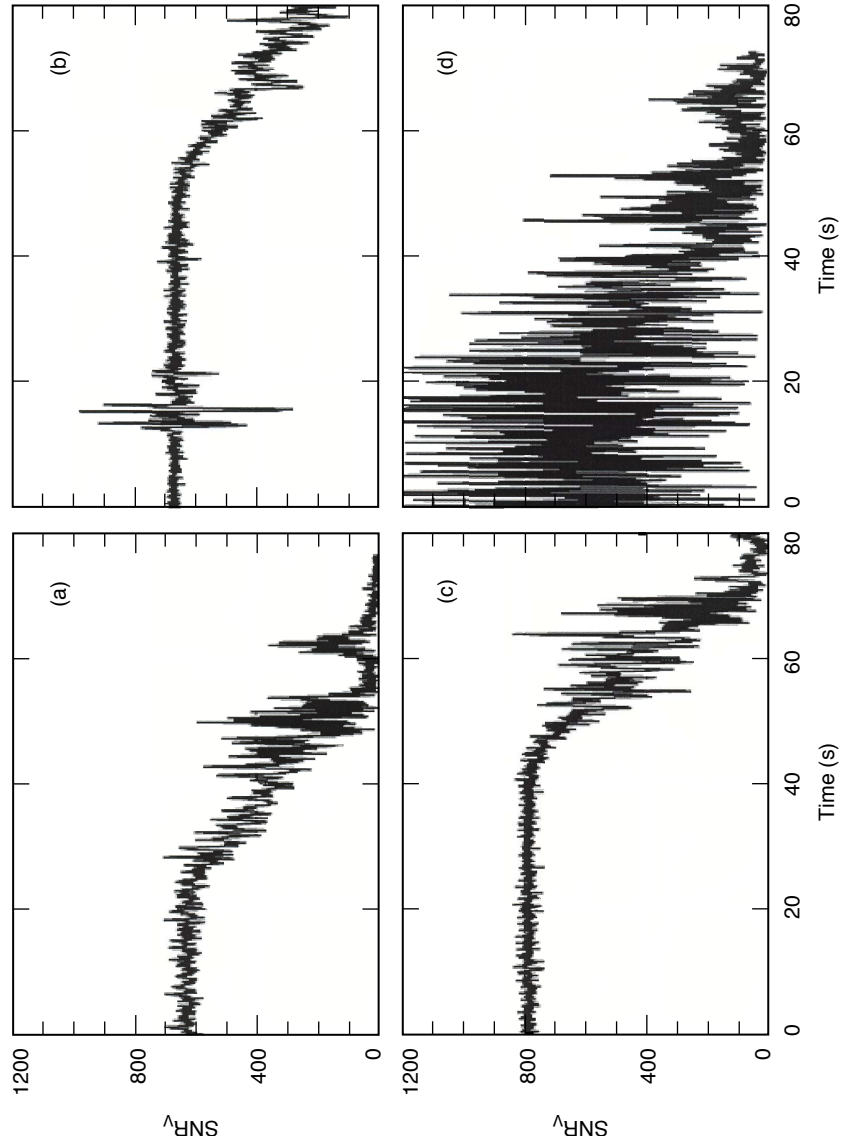


Fig. 1-8. Voltage SNR profiles from four CHAMP occultations in October 2001: (a) probable ducting in the lower troposphere (at ~54 to 62 s), (b) sporadic E-layer (at ~15 s), (c) strong, clean signal above the tropopause (at ~55 s), and (d) troublesome.

Figure 1-9 shows a phasor diagram in the complex plane containing a snapshot at a given instant of the phasors for three interfering rays, the “main” ray m , and the additional rays a and b . This is one of the simplest multipath scenarios, which usually comes in odd-numbered packets, 3, 5, 7, ..., except at caustic contact points, or when a reflected ray is present. The vector sum

$$\hat{E}(t) = E_m(t) + E_a(t) + E_b(t) \quad (1.4-1)$$

gives the observed amplitude and excess phase. The excess phase and amplitude of the individual components are not directly observable. The amplitudes of these rays, E_m , E_a , and E_b , usually differ and they also vary with time, but usually at a much slower rate than $\hat{E}(t)$ can vary because of interference. Their amplitudes are determined by their respective defocusing factors, which usually vary relatively slowly except near a caustic contact point. The individual excess phase of each contributing ray, ϕ_m , ϕ_a , or ϕ_b , depends on its individual ray path through the medium, which is changing with time as the ray path tangency point migrates downward or upward through the medium as a result of the orbital motions of the satellites. Because there are about 10^8 wavelengths along the ray path, the phases of these rays change rapidly with time. The relative phases, $\phi_a - \phi_m$ and $\phi_b - \phi_m$, vary at a rate that depends on the size of the refractivity perturbation that caused the multipath and on the local defocusing by smooth dry air. Typically this difference in rate is below 10 to 15 Hz. The relative values of the amplitudes E_a and E_b , compared to the amplitude of the main ray E_m , determine whether we have deep scintillation or light interference from these particular rays. Each of these three rays has its own excess Doppler signature because the directions of their ray path tangent vectors leaving the transmitting GPS satellite and arriving at the LEO are different. The excess Doppler relationship given in Eq. (1.2-3) is valid for each of these contributing rays, but not to the Doppler for $\hat{E}(t)$. In

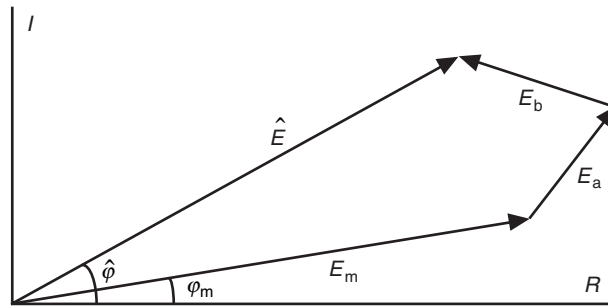


Fig. 1-9. Complex diagram showing composite amplitude \hat{E} and phase $\hat{\phi}$ from three multipath rays.

other words, $d\hat{\phi}/dt$ does not yield per se a bending angle from Eq. (1.2-3) because $\hat{E}(t)$ does not represent an actual ray, only a composite.

The interesting point about these multiple rays is that, although they arrive at the LEO at the same time, they all have different excess Doppler values because their tangent vectors at the LEO differ. Thus, from Eq. (1.2-3) it follows that when we have multiple rays arriving concurrently at the LEO, they arrive with distinct bending angles.

Figure 1-10 shows a schematic diagram for a hypothetical multipath scenario in which up to 5 rays concurrently arrive at the LEO. The ordinate is the excess Doppler shift for individual rays, and the abscissa is the epoch of the observation at the LEO. For this scenario, the phase and amplitude contributions from possibly 1, 3, or 5 rays will simultaneously be registered in the total amplitude and phase measured at the LEO. The actual number of contributing rays depends on the observation epoch t_k , $k = 1, 2, \dots$, as shown in Fig. 1-10(a). The measured composite phase will be an unknown combination of the phases of the individual rays weighted by their respective amplitudes per the vector diagram in Fig. 1-9.

1.4.1 Spectral/Holographic Techniques

What this multipath problem, described in Fig. 1-10, needs is a transformation that converts the multi-valued time series of observations (in Doppler or bending angle) into a single-valued series. There are several ways of accomplishing this. One is to convert the multi-valued time series into a single-valued spectral series, as suggested in Fig. 1-10(b). Here ω is a spectral variable, for example, a Fourier variable from a fast Fourier transform (FFT)

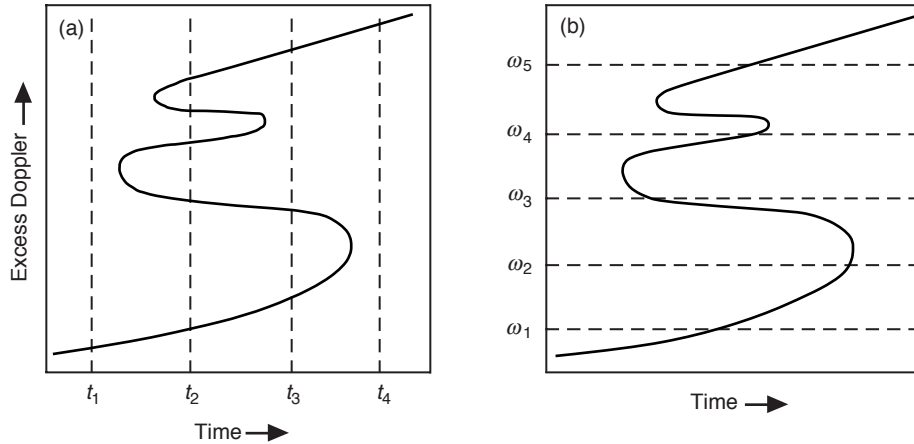


Fig. 1-10. Hypothetical multipath scenario for a setting occultation: (a) the phase at the LEO is measured at the epochs t_k , $k = 1, 2, \dots$, resulting in multiple Doppler contributions at some epochs, and (b) the time series of observations is transformed into a single-valued spectral series.

that can be made proportional to or at least related in a one-to-one way to the excess Doppler. In Fig. 1-10(b), we have converted through a Fourier transform the time series of phase and amplitude measurements of the field, in which the bending angle may not be a unique function of time, into a spectral series in which the bending angle is a unique function of the spectral variable. The peak power in the Fourier spectrum locates the excess Doppler from an individual ray and, therefore, the bending angle per Eq. (1.2-3). This approach was applied to the open-loop phase and amplitude measurements observed from the Earth from the occultation of Voyager 2 by Uranus in 1986 [10]. Figure 1-11 shows an open-loop power spectrum over time, composed of contiguous strips of 10-s temporal width and an excess Doppler breadth of about 15 Hz.

The word “radio-holography” is often used to describe this class of “wave/optic” analysis techniques. Radio-holographic techniques first “stop” the phase rate of the LEO observations by subtracting a time-dependent phase predicted from a realistic model that includes both geometric delay from satellite kinematics and the refractive delay [62–69]. This narrows the effective bandwidth around the spectral peak of the reference ray used in the reference model. The variability of the Doppler shift in the model typically ranges from several hertz to at most a few tens of hertz. Then a complex spectral algorithm,

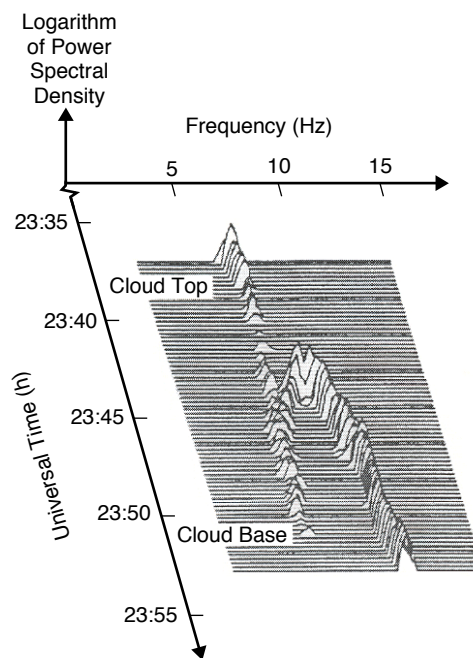


Fig. 1-11. The power spectrum of the radio signal from Voyager 2 while occulting behind Uranus' lower troposphere in 1986. Redrawn from [10].

for example, an FFT, is applied to obtain the power spectrum over the relatively narrow remaining bandwidth of the stopped phase profile within a prescribed time window or time width. Maximum power yields the Doppler tone, or tones when atmospheric multipath is present. From Eq. (1.2-3), one then obtains the bending angle for each tone. The temporal evolution of the tones is obtained by sliding the sample time window used in the FFT. Under the assumption of spherical symmetry, the one-to-one relationship between the recovered excess Doppler/bending angle and the impact parameter of the ray obtained from Bouguer's law yields the value of the impact parameter. From this approach, one recovers a unique bending-angle profile versus impact parameter for a given ray member, but corrupted as always by measurement noise and modeling errors. From this point, the Abel transform yields the refractivity profile versus impact parameter, and then from the relationship $a = r_* n(r_*)$ at a tangency point, one obtains $n(r)$.

Figure 1-12 shows two snapshots of the bending-angle spectra recovered from a GPS/MET occultation over the Sea of Okhotsk in the Russian Far East using a holographic technique [29,54,67,68]. Figure 1-12(a) shows a narrow (half-power width $\sim 20 \mu\text{rad}$) single tone in the mesosphere at 56-km altitude. This translates into a vertical resolution of about 60 m. Figure 1-12(b) shows the complex tone structure about 1/2 minute later for the same occultation. Here the tangency point of the direct ray lies deep in the lower troposphere between 1 and 2 km above sea level. This figure also includes a weaker tone corresponding to a near-specular reflection from the ocean.

1.4.2 Back Propagation

A second, fundamentally different approach to achieving an equivalent single-valued time series and to improving resolution is to map the field measurements recorded by the LEO onto another surface much closer to the refracting medium. This technique is based on the Helmholtz–Kirchoff integral theorem from classical electrodynamics. This integral theorem expresses the amplitude and phase of an electromagnetic wave at a given point in terms of an integral involving the distribution of the amplitude and phase of the wave over an enclosing surface. It explicitly accounts for the retardation time between any point on the radiating surface and the interior point, that is, the travel time between these points resulting from the finite speed of light. The theorem is valid when the scale of the radiating surface is very much larger than the wavelength of the wave. Both the Rayleigh–Sommerfeld and Fresnel–Kirchoff scalar diffraction theories follow directly from this asymptotic theorem [48,70]. These scalar diffraction theories enable both the forward-propagation techniques using a phase screen model to mimic the observations [71–74] and the backward-propagation techniques to map the observations to a back plane [75–77].

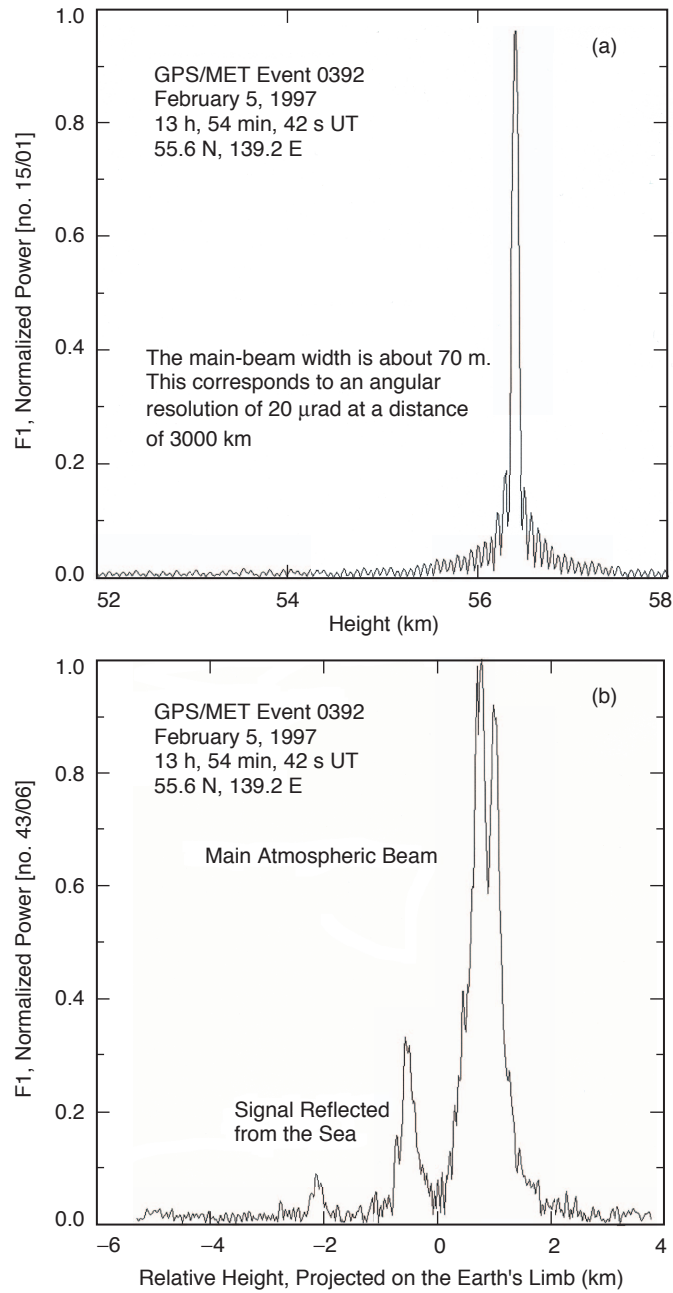


Fig. 1-12. Bending-angle spectra from a GPS/MET occultation using the radio holographic technique [67]: (a) narrow spectral distribution from a single ray in the mesosphere and (b) complex ray structure low in the troposphere, including an ocean surface reflection.

The integral theorem also can be used to propagate a wave through a succession of surfaces. For example, one form of the multiple phase screen technique involves placing successive parallel planes spaced along the paraxial direction of a wave through an inhomogeneous medium. The planes are mounted perpendicular to the paraxial direction, essentially the preferred direction of propagation. The medium between the screens is taken as homogeneous. Therefore, to compensate for the actual inhomogeneous medium between any two adjacent screens, a phase offset is added to the wave in the subsequent screen. This technique can be more tractable for propagating a wave than solving Maxwell's equations. In two-dimensional problems, the multiple phase screen technique involves a one-dimensional scalar diffraction integral applied successively, whereas Maxwell's equations form a second-order system that in general requires simultaneous integration over two dimensions [78–82].

In the multipath problem, the mapped or back-propagated field (toward the emitter) can provide an equivalent virtual time series of phase and amplitude values in another surface that has the favorable property of no or at least substantially fewer multi-valued virtual Doppler points. The one-dimensional back-propagation diffraction integral in a vacuum (see Eq. (A-22) in Appendix A) is given by

$$E(\mathbf{r}_1) = \sqrt{\frac{i}{\lambda}} \int_C \left(\frac{E(\mathbf{r}_2)}{\sqrt{r_{12}}} \exp(-ikr_{12}) (\hat{\mathbf{r}}_{12} \cdot \hat{\mathbf{n}}(\mathbf{r}_2)) \right) ds_2 \quad (1.4-2)$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$; \mathbf{r}_2 denotes a point at the LEO; ds_2 denotes an incremental arc length along the path C , defined by the trajectory of the LEO over which observations were made during the occultation; and $\hat{\mathbf{n}}(\mathbf{r}_2)$ is the outward unit normal vector to C . The intervening medium is a vacuum. We apply this path integral using the LEO observations of phase and amplitude to obtain the mapped field at the position \mathbf{r}_1 . Because maintaining phase coherence in this diffraction integral is so important, the three-dimensional relative motion of the transmitting GPS satellite and the LEO over the integration span of the occultation observations must be accurately modeled. Also, phase connection in the mapped phase along the back surface must be maintained. In addition, the stationary-phase technique can be used to set practical integration limits in the diffraction integral for the LEO phase and amplitude measurements as a function of the position \mathbf{r}_1 . These and other details are found in [76,77,83].

Figure 1-13 provides a one-dimensional schematic for the concept. Here the LEO travels vertically downward in the LEO plane and vertical distance in this plane is proportional to elapsed time, 2 to 3 km/s. Multiple rays from different altitudes in the atmosphere arrive concurrently at the LEO plane. We use the scalar diffraction integral in Eq. (1.4-2) to propagate the field backward from

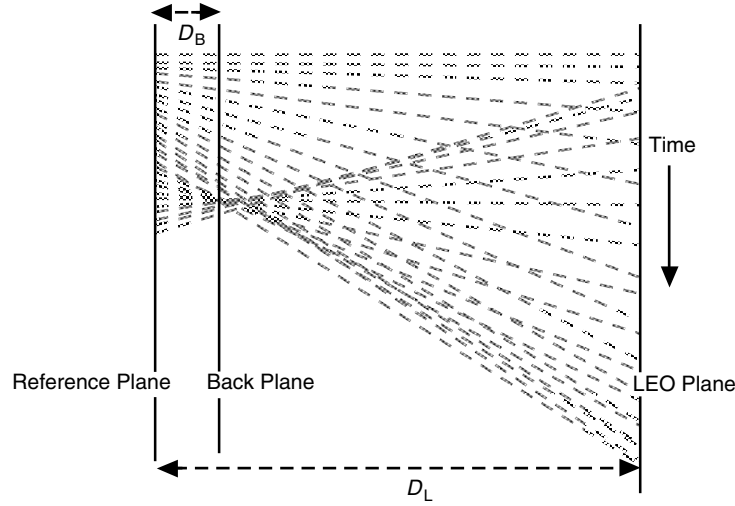


Fig. 1-13. Back propagation geometry. The measured field in the LEO plane is back propagated to the back plane to reduce multipath.

the LEO plane toward the emitter to the position of the back plane mounted perpendicular to the LEO/GPS line. Figure 1-13 suggests that a significant reduction in multiple rays can be achieved from transforming the observations to an equivalent “more focused” set in a back plane [76,77,83], particularly when the actual observation distance is large compared to the vertical scale in the atmosphere over which multipath occurs. The limb distance of a LEO is about 3000 km, but the altitude range in the troposphere from which most multiple rays arrive at the LEO is 10 km or less. Reducing this at least 300 to 1 ratio to 30 to 1 or even 10 to 1 can improve the multipath problem and reduce the Fresnel zone.

Here is a heuristic way of looking at the back-propagation concept, which is further discussed in Chapter 2. To simplify the math, let us place the emitting GPS satellite at infinity, as indicated in Fig. 1-4, and we assume a circular orbit for the LEO. Consider a point (r, θ) well out of the atmosphere ($n \equiv 1$) through which a ray passes after traversing the atmosphere. Suppose we displace that point by a small angular increment to $(r, \theta + \Delta\theta)$, holding the radius r fixed. A new ray arrives at that new point with a new impact parameter value, $a + \Delta a$. Then the difference between the excess phase on the new ray and the excess phase on the old one through the point (r, θ) is given by $\Delta\phi = (\partial\phi / \partial\theta)\Delta\theta$. From Eq. (1.2-3), we have

$$\frac{\partial\phi}{\partial\theta} = k \left(T \cdot \frac{\partial r}{\partial\theta} - \frac{\partial r_{GL}}{\partial\theta} \right) = -k(a - b) = -k \left(D\alpha + a \frac{\alpha^2}{2} + \dots \right) \quad (1.4-3)$$

where $b = |\mathbf{r} \times \mathbf{r}_{\text{GL}}| / r_{\text{GL}} \rightarrow r \sin \theta$, the impact parameter of the straight line between points G (now placed at infinity along the direction $\theta_G = \pi$ in Fig. 1-6) and L. Also, from Eq. (1.2-4) we have Bouguer's law for the impact parameter of the ray, $a = |\mathbf{T} \times \mathbf{r}| = r \sin(\theta + \alpha)$. For collimated incident rays, the angle γ between the ray and the radius vector is simply $\gamma = \theta + \alpha$. These quantities are indicated in Fig. 1-4. The distance D is given by

$$D = r \cos(\theta + \alpha) \quad (1.4-4)$$

D is essentially the limb distance minus the small increment $a\alpha$. It is the distance between the point (r, θ) and the point (a, θ_a) with $\theta_a = \pi/2 - \alpha(a)$, as shown in Fig. 1-4. The latter point is on the impact parameter space curve associated with the rays after their encounter with the atmosphere. The excess Doppler shown in Fig. 1-10 is proportional to $\partial\phi/\partial\theta$ in Eq. (1.4-3). Multiplying Eq. (1.4-3) by $\lambda\dot{\theta}_L/2\pi$ yields the relationship between excess Doppler (assuming that $\dot{r}_L = 0$) and the bending angle

$$\lambda f_D = -D\dot{\theta}_L\alpha + O[\alpha^2] \quad (1.4-5)$$

where $\dot{\theta}_L$ is the projection of the LEO orbital rate in the plane of propagation, typically in the range 0.7 to 1.0 mrad/s.

We now construct $\partial^2\phi/\partial\theta^2$. When θ is varied and r is held fixed, $\Delta a = r \cos(\theta + \alpha)(\Delta\theta + \alpha'\Delta\theta)$, where $\alpha' = d\alpha/da$. Therefore,¹

$$\left. \begin{aligned} \frac{\partial a}{\partial \theta} &= \frac{r \cos(\theta + \alpha)}{1 - \alpha' r \cos(\theta + \alpha)} \\ \frac{\partial b}{\partial \theta} &= r \cos \theta \end{aligned} \right\} \quad (1.4-6)$$

It follows from Eqs. (1.4-3) and (1.4-6) that

$$\frac{\partial^2\phi}{\partial\theta^2} = -k \left(\frac{D^2\alpha'}{1 - D\alpha'} + a\alpha + O[\alpha^2] \right), \quad D = r \cos(\theta + \alpha) \quad (1.4-7)$$

¹ Incidentally, Eq. (1.4-5) provides an expression for estimating the duration of an occultation in the neutral atmosphere. Noting that $\Delta\theta = \int (\partial a / \partial \theta)^{-1} da$, we have

$\Delta T = \dot{\theta}_L^{-1} \int_{a_1}^{a_2} (\partial a / \partial \theta)^{-1} da \doteq (a_2 - a_1 + D(\alpha_1 - \alpha_2)) (\dot{\theta}_L D)^{-1} \approx 50 - 100 \text{ s}$. The duration strongly depends on the magnitudes of $\dot{\theta}_L$ and α near sea level.

The first point to notice is that, on a highly defocused ray with $|D\alpha'| \gg 1$, it follows that $\partial^2 \varphi / \partial \theta^2 \rightarrow kr \cos \theta$. Therefore, the LEO excess Doppler rate approaches essentially a constant in strong defocusing, that is, $df_D / dt \doteq (\partial^2 \varphi / \partial \theta^2) \dot{\theta}_L^2 / 2\pi \rightarrow \dot{\theta}_L^2 r \cos \theta_L / \lambda$, or 10 to 15 Hz/s, depending primarily on the value of $\dot{\theta}_L$. This is an important coherence issue for holographic techniques applied to the lower troposphere. Holographic techniques subtract a model phase from the observations to “stop” their phase variability, i.e., to greatly narrow their spectral bandwidth, making an FFT or a similar complex spectral analysis practicable. The near-constancy of the excess phase acceleration in highly defocused areas is an important factor in the accuracy of the modeled phase profile.

The second point about Eq. (1.4-7) concerns choosing a value for the back-plane distance D_B . Referring to the excess Doppler profile shown in Fig. 1-10 evaluated along the LEO orbit, the undulations in the phase rate curve in the back plane will be greatly diminished if we choose $D_B \ll D$. Basically we want $\partial \varphi / \partial \theta$ to be monotonic, no turning points with respect to θ , or with respect to time, or with respect to any other displacement metric, for example, arc length along the path defined by the intersection of the back plane with the plane of propagation. It follows from Eq. (1.4-7) that this is equivalent to requiring that the defocusing on the back plane be such that $\zeta_B^{-1} = 1 - D_B(d\alpha/da) > 0$ for all values of a spanned by the occultation episode. This condition on the value of back-plane distance D_B , i.e., that $D_B(d\alpha/da) < 1$ for any point along the back plane, ensures monotonicity in excess phase rate versus displacement along the back plane if we can place the back plane close enough to the impact parameter space curve defined by (a, θ_a) . This may not be achievable with a fixed back plane.

Why not set $D_B = 0$? No multipath and no defocusing there. It turns out that that is exactly the right choice, but the problem is that one knows a priori the value of neither α nor a . From Eq. (1.4-4), it follows that setting $D_B = 0$ is equivalent to placing the back plane at the point (a, θ_a) with $\theta_a = \pi/2 - \alpha(a)$. Setting $D_B = 0$ is impossible to uniformly accomplish for all values of a with a fixed back plane because the impact parameter space curve as a function of a is in general not straight, and it is markedly non-linear in multipath zones.

So, with the back-plane methodology, some compromise in choosing D_B must be made. The difference in phase between the point (a, θ_a) on the impact parameter space curve and a nearby point (r_B, θ_B) on the back plane is given to first order by

$$\Delta \varphi = -a \Delta \theta + 0 \Delta r \quad (1.4-8)$$

where $\Delta\theta = \theta_B - \theta_a$ and $\Delta r = r_B - a$. The first-order phase change in the radial direction is zero because the ray through the point (a, θ_a) is orthogonal to r_a at that point. We select $r = r_B$ and $\theta = \theta_B = \pi/2 - \alpha_B$ as the fixed point to position the back plane perpendicular to the LEO/GPS line, with α_B given a fixed value. Then the phase difference $\Delta\phi$ for a given value of a between the point (a, θ_a) on the impact parameter space curve and a nearby point on the back plane is given to first order by $\Delta\phi = -a(\theta_B - \theta_a) + 0(r_B - r_a) = -a(\alpha(a) - \alpha_B)$. We now have a new caustic possibility in the back plane, $d\phi/da = \alpha_B - \alpha - a d\alpha/da = 0$. This arises on the back plane whenever it is away from the impact parameter space curve $r = a$, $\theta_a = \pi/2 - \alpha(a)$, generated by varying a . Since most caustics occur where refractivity gradients are large, and therefore where α is large, a good rule of thumb to reduce the probability of caustics occurring on the back plane has been to set $\alpha_B = \alpha_{\text{Max}}$, where α_{Max} is the largest typical bending angle encountered in the lower troposphere, where major multipath is likely to occur, 30 to 40 mrad. This choice places the back plane close to the impact parameter space curve in that vicinity [65,83]. Even though this plane is in the Earth's atmosphere, it is treated as though it were in a vacuum. It serves only as a platform for recovery of the refractivity profile.

We designate the mapped phase using the diffraction integral in Eq. (1.4-2) from the LEO to the back plane as $\phi_B(s_B)$, where s_B is path length along the back plane perpendicular to the LEO/GPS line. From Eq. (1.2-3) or Eq. (1.4-5), it follows that, when $d\phi_B/da$ is monotonic in the back plane, then the bending angle of a ray intersecting the back plane can be inferred unambiguously from the directional derivative of $\phi_B(s_B)$. For collimated incident waves, it follows that

$$\frac{d\phi_B}{ds_B} = -k \sin \alpha \quad (1.4-9)$$

By assuming spherical symmetry, one obtains the impact parameter for each bending angle using Bouguer's law. From the profile of bending-angle and impact parameter pairs, one can use the standard geometric optics Abel transform technique to recover the refractivity profile. In severe multipath situations, the caustic condition on the back plane can be violated, but experience with actual observations has established that the technique significantly mitigates multipath problems [83].

The back-plane technique gains two benefits, reduced multipath and also enhanced vertical resolution. The Fresnel zone is smaller at the back plane by roughly the factor $\sqrt{|D_B \zeta_B / D_L \zeta_L|}$.

1.4.3 The Canonical Transform Technique

We have noted that rays are stationary-phase paths. Therefore, their space coordinates can be transformed with Hamilton–Jacobi theory into canonical coordinates or ray coordinates. The canonical transform technique [84] essentially uses Hamilton–Jacobi theory to transform the two-dimensional space coordinates (z, x) and their slopes dx/dz for a ray system, that is, for a family of rays generated by varying, for example, the impact parameter a , into three-dimensional phase–space coordinates (z, x, a) . Here x is vertical directed in Fig. 1-4 and z is horizontal directed. In these transformed coordinates, a becomes the canonical conjugate momentum to the variable x . One then applies a Fourier integral transform similar to the scalar diffraction integral in Eq. (1.4-2) to map the field from the phase and amplitude measurements along the LEO path in space coordinates onto the a -plane in phase–space coordinates, from which the phase $\varphi(a)$ is obtained. On this a -plane, $k\alpha = -d\varphi(a)/da$. The technique yields an unambiguous determination of α and a , provided that spherical symmetry holds and super-refractive zones are avoided, and subject, of course, to the quality of the LEO observations and the modeling. The technique fails in super-refractivity zones because it is in part ray theory. The technique avoids the potential back-plane caustic and multipath problems of the general back-propagation technique, and its vertical resolution potential is comparable to wave theory.

1.4.4 The Impact Parameter Space Curve

Another less elegant but conceptually straightforward method for avoiding caustic and multipath problems maps the phase and amplitude measurements made by the LEO along its path to the impact parameter space curve associated with the rays after their encounter with the atmosphere. As Fig. 1-4 indicates, the rays are ordered through a transverse plane located perpendicular to the incoming rays at a position prior to their entering the atmosphere. This pre-encounter impact parameter space curve is a simple straight line perpendicular to the incoming collimated rays, or a circular arc of radius $\rho_G/2$ if the incoming rays are from a spherical wave. This is equivalent to the proposition that the rays are single-valued as a function of impact parameter a before encountering the atmosphere; a one-to-one relationship holds between the value of the impact parameter of the ray intersecting the transverse plane (or arc) and the position of that intersection point on the plane (or arc). If spherical symmetry holds, the rays will continue to be single-valued after passing through the atmosphere, that is, each ray will continue to have a unique value of the impact parameter. Therefore, after transecting a spherical symmetric atmosphere, there will remain a one-to-one relationship between the impact parameter of a ray and the angle through which it was refracted. The point

$r_a = a$, $\theta_a = \pi/2 - \alpha(a)$, located on an outgoing ray, also preserves this one-to-one relationship with that ray, if spherical symmetry holds. Bouguer's law in Eq. (1.2-4) requires it. The curve traced out by the point (r_a, θ_a) as a is varied is the space curve for the impact parameter *after* atmospheric encounter. As mentioned earlier, there will be no caustics on this post-encounter curve, but it is no longer a simple curve because of the dispersive refractive gradient.

It is important to keep in mind that we are not so much interested in ray reconstruction or mapping the electromagnetic field through space as we are in recovering the refractivity profile. Any construction, however non-physical, that facilitates refractivity recovery becomes a valid candidate for this application. In this regard, a fictitious curve placed after a ray's atmospheric encounter that denumerates the impact parameter uniquely along its path should suffice for spherical symmetric geometry because the rays through points in this curve will be unique [77]. This holds even though the impact parameter space curve actually lies within the atmosphere.

On the post-encounter impact parameter space curve, the phase of the intersecting ray with an impact value a can be obtained from the defining integral for phase delay given in Eq. (1.2-2) using Bouguer's law. Let this phase be defined by $\varphi_a(a, \theta_a)$ [with $\theta_a = \pi/2 - \alpha(a)$]. It can be shown from Eq. (1.2-2) (with the GPS transmitting satellite at infinity and the phase referred to the line $\theta = \pi/2$) that

$$\varphi_a(a, \theta_a) = k \left(a\alpha(a) + \int_a^\infty \alpha(\omega) d\omega \right) \quad (1.4-10)$$

The space curve itself is generated by the tip of the vector $\mathbf{r}_a(a)$, with $r_a = a$ and $\mathbf{r}_a \cdot \mathbf{T}_a = 0$, where \mathbf{T}_a is the unit tangent vector of the ray passing through the tip of \mathbf{r}_a . The angle ψ_a between the tangent vector $d\mathbf{s}_a$ to the impact parameter space curve and \mathbf{r}_a is given by

$$\left. \begin{aligned} \cos \psi_a &= \frac{\mathbf{T}_a \cdot d\mathbf{s}_a}{ds_a} = \frac{1}{\sqrt{1 + (a\alpha')^2}} \\ \sin \psi_a &= \frac{-a\alpha'}{\sqrt{1 + (a\alpha')^2}} \end{aligned} \right\} \quad (1.4-11)$$

with

$$\frac{ds_a}{da} = \sqrt{1 + (a\alpha')^2} \quad (1.4-12)$$

From Eq. (1.4-10), it follows that

$$ka \frac{d\alpha}{da} = \frac{d\varphi_a}{da} \quad (1.4-13a)$$

or

$$\alpha(a) = -\frac{1}{k} \int_a^\infty \frac{1}{a} \frac{d\varphi_a}{da} da \quad (1.4-13b)$$

By replacing φ_a in this integral with the mapped phase $\varphi_a(\mathbf{r}_a)$ obtained from the back-propagating diffraction integral of the LEO observations given in Eq. (1.4-2), we recover the profile for $\alpha(a)$. Note that, from Eq. (1.4-10), $\varphi_a \rightarrow 0$ as $a \rightarrow \infty$ if $\alpha \rightarrow 0$.

Chapter 2 discusses one method for recovering the impact parameter space curve $(a, \theta_a(a))$ starting from an initial known point $(a_1, \theta_a(a_1))$ on the curve. It uses the back-propagation diffraction integral to map by iteration the LEO observations to a converged point $(a_1 + \Delta a, \theta_a(a_1) + \Delta \theta_a)$ near the initial point. Each successive converged point on that curve yields the recovered values for $\varphi_a(a_k)$ and $\alpha(a_k)$, $k=1, 2, \dots$. For a spherical symmetric atmosphere, the recovered impact parameter space curve can be used with the Abel transform to recover $n(a)$. This method also fails in super-refractive layers. Also, Eq. (1.4-13) shows that the actual shape of the (a, θ_a) curve, however complicated it might be, is of no theoretical consequence in recovering $\alpha(a)$, although a complicated shape could slow the iteration process.

The difference between the phase $\varphi(a)$ on the a -plane in phase space, which uses the canonical transform technique, and $\varphi_a(a)$ on the post-encounter curve in space coordinates (a, θ_a) is simply

$$\varphi(a) = \varphi_a(a) - ka\alpha(a) = k \int_a^\infty \alpha da \quad (1.4-14)$$

1.4.5 A Full-Spectrum Wave Theory

The various spectral or holographic approaches and the scalar diffraction approaches, both the forward- and backward-propagation techniques, involve a hybrid of wave and ray theory concepts. In the end, profiles of bending angle versus impact parameter are obtained from which the refractivity is recovered using the Abel integral transform.

In the quest to deal with multipath problems and related resolution issues, there is yet another approach that seems not to have been fully considered in radio occultations. This involves a full-spectrum wave-theoretic treatment of the electromagnetic propagation through a refracting medium. The technique presented here uses a modified Mie scattering approach adapted to an

inhomogeneous medium. It is applicable to the more general case of a scattering surface embedded in a transparent, spherical refracting medium. An interesting feature of this approach is its full-spectral series representation of the wave.

The spectral series is a solution to the Helmholtz equation for the time-independent part of the wave. Each spectral component of integer number l consists of a variable spectral coefficient multiplied by the basis functions, which in their native form are the spherical Hankel functions of integer order l and the spherical harmonic functions. The effect of the inhomogeneity in the propagation medium is contained in the spectral coefficients, leaving the basis functions unchanged from their functional form in a homogeneous medium.

To derive the variation of the spectral coefficients across a spherical surface, one invokes the continuity conditions from electrodynamics that apply to the radial and tangential components of the incident, reflected, and transmitted electromagnetic field vectors, just as is done in Mie scattering theory. A limiting process yields the differential equations describing the variation of the spectral coefficients in a stratified medium. The accuracy of the technique deteriorates when the truncation assumption about the smallness of certain curvature terms in the field equations becomes invalid, for example, near a turning point. The related Wentzel–Kramers–Brillouin (WKB) method has similar problems at a turning point. The technique is not applicable to a medium with significant back scattering, but it is adaptable to an absorbing medium.

In this wave theory, each spectral coefficient contains a spectral density function for the phase delay induced by the refractivity gradient of the atmosphere. Basically the spectral density function gives the cumulative phase delay induced by the atmospheric refractivity gradient on the l th spectral component of a wave that has traveled from outside the atmosphere down to a given radial distance. This spectral density for phase delay, which is a function of the radial coordinate and the spectral number, can be formally written in terms of an integral over the radial coordinate that explicitly contains the refractivity gradient. Without a refractivity gradient, there is no phase delay in the spectral coefficient.

This series representation of the wave can be directly linked to the Fourier transform of the time series of LEO-observed amplitude and phase. The spectral density function can be recovered by taking the Fourier transform of the stopped version of the formal spectral series from wave theory and the Fourier transform of the stopped phase and amplitude observations made by the LEO over time. Equating these two transforms enables recovery of the spectral density function for the phase delay. The refractivity gradient follows from an inversion of the defining integral for the derivative of the spectral density function.

In a wave theory, the fundamental observations are the phase and amplitude measurements of the field, not Doppler observations or the resolving of possible ambiguities contained therein from multipath. There is no need to propagate the field or map it from one surface to another. The spectral series gives the field at an arbitrary point, even inside of the refracting medium, provided turning points are avoided and the curvature truncation assumption is valid. Multipath and caustic rays become almost irrelevant in wave theories. This applies to parabolic equation techniques with multiple phase screens and to the full-spectrum technique described here. A wave theory calculates the phase and amplitude of a caustic ray as accurately as any other ray. The predicted amplitude of the field at the LEO when it is located on a caustic ray is proportional to $|d^2a/d\theta^2|^{1/3}$ in wave theory, unlike the infinite value predicted in (second-order) geometric optics.

In a full-spectrum wave theory, there will be stationary-phase points in spectral number from where the principal contributions to the summation of the spectral series originate. These stationary-phase points in spectral number closely correspond to the impact parameter values (in phase units) of the multiple rays, including caustic contact points. But in the full-spectrum technique, all points in spectral number are swept up together into a spectral density function that holds for all relevant spectral numbers, not just stationary-phase points.

As mentioned above, we are more interested in the refractivity recovery than in calculating the electromagnetic field vector at the LEO. In this regard, the main emphasis in this full-spectrum technique is on computing the spectral density function for the phase delay, a somewhat simpler task than computing the electromagnetic wave from the spectral series itself. Moreover, being fundamentally a result from wave theory, the recovered refractivity profile has the potential for a much finer radial resolution than dictated by the local first Fresnel zone. This also applies to certain holographic and scalar diffraction techniques mentioned above. Theoretical vertical resolution for wave techniques is proportional to the carrier wavelength divided by the angular synthetic aperture of the observations $\dot{\theta}\Delta T$, or roughly $150(\Delta T)^{-1}$ m for a LEO. Here ΔT is in seconds; it is the minimum of either the time span of the selected observation sequence or the coherence limit of the observations. This would suggest a dekameter-level resolution potential, but the actual vertical resolution will be limited by atmospheric vicissitude, such as horizontal variations in refractivity caused by advection and turbulence, and also by other GPS signal structure and data processing limitations [72]. Perhaps 10 percent of the vertical axis of the local first Fresnel zone is a practicable goal, although the evidence for this conjecture is still inchoate at this time.

Full-spectrum wave-theoretic approaches are powerful for a certain class of geometries, but they are notoriously cumbersome. That is why so many

alternate schemes for propagating waves through complex media have evolved over the years since Mie scattering theory was first formulated in the early 20th century. The use of the Helmholtz–Kirkchoff integral theorem for multiple phase screen propagation is one example. Other parabolic wave equation techniques for propagating waves through a medium can be found in [81]. Spherical symmetry or some similar symmetry is almost *de rigueur* by definition in a full-spectrum wave-theoretic approach. When the wavelength is small compared to the scale of the refracting medium, the spectral series are slow to converge. The basis functions in the spectral representation of the wave, the spherical Bessel functions, and the spherical harmonic functions become difficult to evaluate in their native forms for large values of spectral number.

Nevertheless, accurate asymptotic forms for the Bessel functions in terms of the Airy functions exist that are quite comprehensive and easy to use [85,86]. Most computer mathematics programs carry a full library of the Airy functions of the first and second kind and their derivatives. Also, the asymptotic forms for the spherical harmonic functions for large spectral number are given by sinusoid functions. While asymptotic functions are used throughout the series, the full-spectral character of the wave equation solution is retained. On the issue of convergence, the stationary-phase technique can be used to identify neighborhoods in spectral number that contribute to the summation of the series representation of the wave, thereby eliminating enormous tracts of non-contributing spectral numbers. Moreover, the stationary-phase technique can be used to establish a duality between certain wave-theoretic quantities evaluated at the stationary-phase points in spectral number and the corresponding quantities in geometric optics. For example, when a stationary-phase point exists, and when super-refractivity situations are avoided, then there is a close correspondence between the spectral density function for phase delay evaluated at a stationary-phase point in spectral number and the eikonal function in geometric optics for the corresponding ray. This is the ray with an impact parameter (in phase units) equal to the stationary-phase point in spectral number.

The essential difference between a full-spectrum wave theory and most holographic techniques [62–69] is in the treatment of the wave itself. Loosely speaking, a full-spectrum wave theory involves a pre-stationary-phase process; holographic techniques involve a post-stationary-phase process. In the full-spectrum technique, the time-independent part of a harmonic wave is described by a spectral distribution within an infinite series summation over spectral number, here an equivalent spectral integral. Holographic treatments describe the wave as a finite sum of the complex amplitudes from the multiple rays involved at an observation epoch, as the schematic in Fig. 1-10 suggests. In the recovery process, both techniques apply a spectral analysis to the phase and amplitude measurements, for example, a fast Fourier transform. The full-spectrum technique recovers the spectral distribution for the phase delay

induced by the refractivity gradient, and it does so with a limit in spectral resolution that is imposed by the uncertainty principle in the Fourier analysis, $\sim 1/\Delta T$, where ΔT is the time interval used in the observations. From the spectral derivative of the recovered spectral distribution for the phase delay, the gradient of the refractivity profile is recovered from inverting the formal integral equation that expresses this spectral derivative in terms of the refractivity gradient. From this one obtains the profile of the change in refractivity over the span of impact parameter values set by the time interval ΔT . One could place, if needed, the recovered spectral density for phase delay into the spectral integral for the wave to compute the stationary-phase points of this integral for successive increments in observation epochs. This gives profiles of the impact parameter values of the multipath rays, including caustic rays if they occur there, the bending angles of the rays, and their angular power spectra.

In comparison, a holographic technique obtains the maximum-likelihood estimate for the bending angle and impact parameter associated with each of the finite number of rays detected in the Fourier analysis of the observations within a given time increment, as well as their angular power spectra. And it then applies the Abel integral transform from ray theory to the recovered bending-angle and impact parameter sequences for each ray to obtain the profile of the change in refractivity over the integration limits. The integration limits in impact parameter for each recovered ray are set to obtain the refractivity profile over the impact parameter range applicable to that ray. Appropriate hand-over at each caustic contact point of the integration limit in the Abel integral from the one nascent ray to the other provides the connections between integral segments, and it yields the entire refractivity profile.

If for certain positions of the LEO there are no stationary-phase values in spectral number in the full-spectrum technique, then in ray theory there will be no rays or eikonal functions to give phase delays. But in a wave theory, there still will be a field predicted for such locations, diminished in amplitude probably, but not zero.

The overviews of Chapters 5 and 6 in Section 1.5 give additional details on this wave-theory approach.

1.5 Overview

1.5.1 Chapter 2

Chapter 2 is largely a revision of an earlier Jet Propulsion Laboratory (JPL) publication [87]. Chapter 2 discusses the changes in the phase and signal amplitude profiles that are observed by a LEO during an occultation of a GPS satellite as a result of a sharp change across a spherical surface in a refraction-related property. This would include (1) the refractivity itself, (2) its scale

height, and (3) the lapse rate of the temperature profile. These sharp changes are embedded in an otherwise smoothly varying refractivity profile. Across the tropopause the lapse rate can change abruptly. Strong refractivity gradients often are associated with a water vapor boundary in the lower troposphere or an electron density change in a sporadic E-layer of the ionosphere. A discontinuous refractivity profile could be a useful limiting case.

Chapter 2 uses a combination of geometric optics and scalar diffraction theory in a thin-screen model to interpret these transients in terms of multipath interference, shadow zones, diffraction, and caustics that are likely to be observed by the LEO. In other words, we primarily address the forward-propagation effects of certain precipitous changes in the vertical profile of the atmospheric or ionospheric refractivity, such as a discontinuous change in refractivity or in one of its derivatives. The applicability of the thin-screen approach in a medium with strong refractive gradients is discussed. The stationary-phase technique for interpretation and for aiding the computation of diffraction integrals is discussed.

This hybrid thin-screen/scalar diffraction approach is used to calculate the phase and amplitude perturbations that would be observed by the LEO from these perturbations in refractivity. Although one customarily thinks of diffraction as producing high-frequency fringe effects, which is true, it also can soften or mitigate the harsh effects predicted by geometric optics, partially filling in the troughs in shadow zones and rounding the peaks in flaring regions where caustics are a factor.

Use of an inverse transform technique to sharpen the resolution of localized features is briefly covered, and it is applied to a specific occultation where the ray tangency point crosses a sporadic E-layer.

1.5.2 Chapter 3

Chapter 3 also is a revision of an earlier JPL publication [88]. To prepare the development of a wave theory in a refracting medium, we review basic Mie scattering theory. Chapter 3 discusses the scattered electromagnetic field from a spherical scattering surface in a homogeneous medium. The field is expressed as series solutions to the Helmholtz equation involving spherical Bessel functions and spherical harmonic functions summed over integer spectral number l . In Mie theory, the scattering coefficients in the spectral series solutions are determined by applying the continuity conditions from electrodynamics that must hold across a boundary bearing a discontinuity ΔN in refractivity. The total scattering from the sphere includes all scattering modes, which arise from external reflection, refraction, and internal reflections.

It is well-known that these spectral series converge slowly with spectral number when the radius r_o of the refracting sphere is very large compared to the wavelength λ of the incident wave. Asymptotic forms for the Bessel and

Legendre polynomial basis functions in terms of the Airy functions and the complex exponential functions, which are applicable when $r_o / \lambda \gg 1$, can be used to greatly facilitate the computation of the scattered field. The asymptotic scattering series is evaluated by converting it into an integration of a phasor multiplied by a slowly varying function of l . Neighborhoods around spectral number points yielding stationary-phase values are the principal contributors to the scattering integrals.

Certain attributes of the scattered field, such as phase, amplitude, reflection, bending angle, defocusing, and caustics, are evaluated at the stationary points in spectral number. A close correspondence is established between these attributes in electrodynamics and their analogues in geometric optics when the observation point is sufficiently away from the shadow boundary.

Rainbow caustics arising from one or more internal reflections within the refracting sphere are evaluated and compared to the predictions from the thin-screen model. For small discontinuities in refractivity, rainbow effects are shown to be prominent only within relatively narrow directions. Third-order stationary-phase theory is used to determine their characteristics in terms of LEO position, the value of ΔN , the ratio r_o / λ , and the number of internal reflections prior to being observed.

Special limiting cases, such as scattering by a perfectly reflecting sphere, by a strongly absorbing sphere, and by a sphere with infinite radius also are discussed in terms of stationary-phase concepts with numerical examples.

When the magnitude of the discontinuity in refractivity at the scattering surface is sufficiently small, $|\Delta N| \sim (2\pi\lambda / r_o)^{2/3}$, about 30×10^{-6} for the Earth's radius or equivalently about 20 mrad of refractive bending at the surface, and when the observation point is sufficiently far from the limb, e.g., at the LEO, then one obtains good agreement in amplitude between the computations based on Mie scattering theory and those based on the thin-screen model combined with scalar diffraction theory.

1.5.3 Chapter 4

In Chapter 4, we review a technique that uses the unitary state transition matrix for the system of first-order electromagnetic wave equations in a refracting medium that is stratified and thin, and whose electromagnetic properties are linear [48,89,90]. This approach has been useful for calculating the propagation of an electromagnetic wave through a thin film. Although not essential to the development of the modified Mie scattering approach, several key propagation concepts are introduced here that are used in the Mie formulation. These include the representation of incoming and outgoing standing waves, the use of osculating parameters, their asymptotic forms, asymptotic matching of incoming and outgoing solutions, and evaluation of the accuracy of the osculating parameter technique.

1.5.4 Chapter 5

In wave theory, we are concerned with solutions to Maxwell's equations. For a harmonic wave in a homogeneous medium, the time-independent parts of the electric and magnetic field vectors are solutions to the Helmholtz equation, which become the components in a spectral series representation of the wave. Each component consists of a constant spectral coefficient multiplied by the basis functions used in the spectral series, which in a spherical medium are the spherical Hankel functions and spherical harmonic functions. Chapter 5 develops a spectral representation for a harmonic electromagnetic wave in a spherical refracting medium. The resulting spectral series for the field in the refracting medium also are expressed in terms of the spherical Hankel and spherical harmonic functions as basis functions. The spectral coefficients for these basis functions are derived for a wave propagating through this refracting medium. The spectral coefficients, which in a homogeneous medium are functions only of spectral number, also vary with radial distance in a refracting medium. Through a modified Mie scattering theory, these spectral coefficients also account for, if applicable, scattering from a spherical surface embedded in the refracting medium. The scattering surface could be defined by a discontinuity in refractivity, or in its gradient or a higher derivative, or by a reflecting surface, or by an absorbing surface.

Chapter 5 uses an osculating parameter technique to calculate the spectral coefficients describing wave propagation in a refracting medium. The method for obtaining the spectral coefficients is similar to a certain parabolic equation technique for wave propagation [81]. In a typical parabolic equation technique, one forms the reduced solution $u(\mathbf{r}) = \exp(-i\mathbf{k} \cdot \mathbf{r})\psi(k\mathbf{r})$ and then eliminates less relevant terms. Here $\psi(k\mathbf{r})$ is a solution to the Helmholtz equation and \mathbf{k} is the wave number vector in the paraxial direction. In our formulation, the form is, for a given spectral number l , $u(r) = \exp(-iG[knr])\psi(knr)$, where $-G[knr]$ is the phase delay incurred by the l th spectral component of the wave *only* from the refractive gradient. The "optical" delay is retained in the original Helmholtz solution $\psi(knr)$. In the continuous medium overlying the scattering surface, the solution is applicable when the medium induces negligible back scattering of the wave, i.e., the medium is transparent. The technique can be adapted to an absorbing medium. It is shown for the limiting case, where the spherical stratification approaches a Cartesian stratification, that the osculating parameter solution becomes the WKB solution for a variable index of refraction. The accuracy of this technique and its range of applicability are addressed. The technique fails at or below a "turning point" (which is spectral-number-dependent in wave theory) because the truncation assumption regarding the smallness of certain curvature terms in the defining differential equations for the electromagnetic field is not valid there. The WKB method also fails for the same reason without appropriate connection formulas between

regimes above a turning point and below, the tunneling regime. On the other hand, one must match the incoming and outgoing spectral coefficients at a turning point to ensure the absence of singularities at the origin from the Hankel functions, and to obtain the correct form for the outgoing spectral coefficients. The outgoing coefficients are the only coefficients applicable at the LEO. Individual Hankel functions are singular at the origin, but the sum of the first and second kinds equals the spherical Bessel function of the first kind, which is well-behaved at the origin. This matching of incoming and outgoing coefficients is accomplished by an asymptotic technique, even though the individual forms for the spectral coefficients fail at and below the turning point.

Chapter 5 is relevant to the applicability of the single thin phase screen/scalar diffraction model, discussed in Chapter 2, to a spherical atmosphere with a significant gradient in refractivity. The thin-screen model serves as a surrogate for the actual atmosphere. Both the thin-screen/scalar diffraction approach and the wave-theory approach lead to predictions of the observed phase and amplitude of the wave having passed through an intervening atmosphere and perhaps having encountered an embedded scattering surface. The wave-theory approach explicitly accounts for scattering and for the effects of the refractive gradient in the surrounding medium on the phase and amplitude of the electromagnetic wave propagating through the medium. In comparison, the thin-screen model lumps these effects into a phase delay profile embedded in the thin screen. This profile mimics the extra phase delay experienced by the wave due to the atmospheric refractivity profile. In accounting for the scattering effects from and through the spherical surface of discontinuity, wave theory also includes the possibility that the scattered wave has undergone one or more reflections inside this surface. Interference, shadowing, diffraction, and caustics can be evaluated using both the thin phase screen/scalar diffraction approach and the full-wave theory approach. The level of agreement between these two approaches and how that level depends on the adversity of the wave propagation conditions in the atmosphere are addressed here. Good agreement with thin-screen results is obtained when “thin atmosphere” conditions hold in a large sphere, i.e., no caustic points in the screen, when rainbow caustic directions are avoided, and when the LEO is some distance from the refracting medium.

Stationary-phase concepts are applied to key observables, such as phase delay, propagation direction, and wave amplitude, the stationary-phase values of which are established in spectral number using a phasor representation in the wave-theory approach. Stationary-phase values for these observables also are obtained in impact parameter space in geometric optics. Chapter 5 establishes a duality between stationary-phase variables in spectral number, when they exist, and their counterparts in geometric optics.

Chapter 5 introduces a general spectral density function $G[\rho, \nu]$ for the phase delay in the spectral coefficient induced by the refractive gradient in the

medium. This function of radial coordinate $\rho = krn(r)$ and spectral number v accounts for the extra phase delay at the radial position ρ on each spectral component induced by the refractive gradient on an incident wave that asymptotically is either planar or spherical at large approaching distances relative to the refracting sphere. Correspondence between this spectral density function for phase delay and key concepts from geometric optics, cumulative bending angle and the eikonal function giving path delay, is established.

Chapter 5 also describes the properties of turning points, caustics, shadow zones, and multipath from a wave-theoretic point of view in a spectral framework. It also discusses these quantities in a second-order geometric optics framework, including its shortcomings near caustics or in dealing with nascent ray pairs with nearly merged impact parameters. Third-order stationary-phase theory is introduced to develop a more accurate ray-theoretic approach near caustics.

Chapter 5 develops a phasor-based approach for evaluating the spectral series using numerical integration combined with the stationary-phase technique. The numerical integration of the spectral representation has been aided by the stationary-phase technique to identify contributing neighborhoods in spectral number, greatly improving their efficiency.

Special topics are addressed—for example, calculating the field when a reflecting surface is embedded in the refracting medium and dealing with turning point computational difficulties where the osculating parameter technique degrades in accuracy. Numerical solutions for the spectral representation of the field at the LEO are presented for various refractivity profiles and scattering surfaces. Numerical examples also include the field at the LEO from refractivity profiles giving multipath, shadow zones, super-refractivity, and caustics, and the field at the LEO from a perfectly reflecting sphere embedded in an overlying refracting medium.

1.5.5 Chapter 6

Here our goal is use the wave-theory concepts in Chapter 5 to develop a technique to recover the refractivity profile from occultation observations of phase and amplitude. The chapter begins with a discussion of GPS receiver operations and ends with a discussion of a recovery technique using the spectral representation developed in Chapter 5. We provide a brief summary here of the recovery sections.

In spherical coordinates, the coplanar components of the time-independent solutions to the Helmholtz equation for the electric field evaluated at the point (ρ, θ) in the plane of propagation may be expressed in spectral series form:

$$\left. \begin{aligned} E_r(\rho, \theta) &= \frac{E_o}{2\rho^2} \sum_{l=1}^{\infty} l(l+1) \left(a_l^-(\rho) \xi_l^-(\rho) + a_l^+(\rho) \xi_l^+(\rho) \right) P_l^1(\cos \theta) \\ E_\theta(\rho, \theta) &= \frac{E_o}{2\rho} \sum_{l=1}^{\infty} \left(a_l^-(\rho) \xi_l'^-(\rho) + a_l^+(\rho) \xi_l'^+(\rho) \right) \frac{\partial}{\partial \theta} P_l^1(\cos \theta) \end{aligned} \right\} \quad (1.5-1)$$

where $a_l^\pm(\rho)$ are the spectral coefficients, $\xi_l^\pm(\rho)$ are the spherical Hankel functions of the first (+) and second (−) kinds, $\rho = krn(r)$, and $P_l^m(\cos \theta)$ are the associated Legendre polynomials with $m=1$. In this coordinate frame, the emitting source is located in the direction $\theta = \pi$. Here E_o is an arbitrary constant when the emitting source is infinitely far, but in the finite case it also accounts for space loss from the wave-front curvature. Wave theory allows for the possibility of two polarization modes for the field vectors, the so-called transverse electric (TE) and transverse magnetic (TM) modes. Here we assume a TM wave, i.e., the magnetic field vector is oriented perpendicular to the plane of propagation. The electric field calculated from Eq. (1.5-1) includes both incoming (−) and outgoing (+) waves, and it takes into account the possibility that radiant energy has come to the point (ρ, θ) via any sector in the refracting sphere, for example, from the bottom limb and from the top limb. But for an occultation geometry from a distant emitting source and when the LEO is located at a point (ρ_L, θ_L) well into the top and outgoing sectors, $\pi/2 - \alpha_L \gg \theta_L \gg 0$, only the terms with plus signs contribute to the evaluation of these series at the LEO when $\rho_L / \lambda \gg 1$.

These spectral series converge only after the value of the spectral number becomes comparable to ρ_o , the radial phase distance of the refracting source, about 2×10^8 for GPS wavelengths. To compute the actual electromagnetic field from these series, one uses special techniques, for example, contour integration in the complex spectral number plane, saddle point integration, and numerical integration aided by stationary-phase theory. One also needs a number of asymptotic forms for the Hankel functions and the Legendre polynomials generally valid for large values of the spectral number l and the radial coordinate ρ . For $l, \rho \gg 1$, the Bessel functions, or spherical Hankel functions in this case, can be written accurately in terms of the Airy functions. Similarly, the spherical harmonic functions can be written in terms of sinusoids.

Hardly any of that computational firepower is necessary when one is interested only in recovery of the refractivity profile. The important aspect for refractivity recovery is the variation in phase of each spectral coefficient $a_l^+(\rho)$, which varies with radial coordinate because of the refractive gradient in the medium. For a refracting medium with negligible back scattering and absorption, these coefficients evaluated at the LEO can be written in the form

$$a_l^+(\rho_L) = f(l) \exp(-i2G^\dagger(v)), \quad v = l + \frac{1}{2} \quad (1.5-2)$$

where $f(l)$ is a function of spectral number. The specific functional form of $f(l)$ depends on the asymptotic boundary conditions for the incoming wave from the emitting GPS satellite. Chapter 5 mainly uses a collimated incident wave for the spectral coefficients, but the adjustments to account for a spherical incident wave are noted there, and they are used in Chapter 6. The quantity $-2G^\dagger(v)$ is the spectral density function of the phase delay for the l th spectral component evaluated at the position of the LEO for a wave that has originated from a distant source; both the LEO and the emitter are assumed to be out of the refracting medium. For a homogeneous medium, $G^\dagger(v) \equiv 0$. Otherwise, $G^\dagger(v)$ is obtained from the general spectral density function $G[\rho, v]$, which is given to sufficient accuracy by the integral

$$\left. \begin{aligned} G[\rho, v] &= \int_{\rho}^{\infty} \left(\frac{d \log n}{d \rho} \right) g(\hat{y}) d\rho, \quad \rho \geq \rho^\dagger(v) = v - \hat{y}^\dagger K_v, \quad g(\hat{y}^\dagger) = 0 \\ g(\hat{y}) &\doteq \pi K_v^2 \left(\text{Ai}'[\hat{y}]^2 + \text{Bi}'[\hat{y}]^2 - \hat{y} \left(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \right) \right), \\ \hat{y} &= \hat{y} \left[\frac{v}{\rho} \right] \doteq \frac{1}{4K_v^4} (v^2 - \rho^2), \quad K_v = \left(\frac{v}{2} \right)^{1/3}, \quad v = l + \frac{1}{2} \end{aligned} \right\} \quad (1.5-3)$$

Here $\text{Ai}[\hat{y}]$ and $\text{Bi}[\hat{y}]$ are the Airy functions of the first and second kind, and \hat{y} is their argument; its exact functional form is given in terms of the ratio v/ρ [85]. The quantity $\hat{y}^\dagger = 0.441 \dots$ is the zero point of $g(\hat{y})$, which is unique. The spatial difference between v and ρ^\dagger is $k^{-1} \hat{y}^\dagger K_{\rho_o} \approx 7$ m, almost negligible. For $\rho < \rho^\dagger(v)$, the form for $g(\hat{y})$ in Eq. (1.5-3) does not apply; the correct form rapidly approaches zero for increasing $\hat{y} > 0$.

The interpretation of the general form $G[\rho, v]$ is as follows: $-G[\rho, v]$ is the phase delay accumulated by the spectral coefficient $a_l^-(\rho)$ for a wave that has traveled from infinity (where $dn/d\rho \rightarrow 0$) down through a refracting medium to a radial distance r with $\rho = krn(r)$. For a given spectral number v , the rate of phase accumulation by the corresponding spectral coefficient $a_l^-(\rho)$ rapidly approaches zero for decreasing ρ values less than v . In this regime, these spectral coefficients are unaffected by the refracting medium overhead or by a scattering surface contained therein. Therefore, $G[\rho, v] \rightarrow \text{constant}$ for decreasing $\rho < v$.

The integration variable $\rho = krn(r)$ is used for convenience. Crossing a super-refractivity boundary where $d\rho/dr$ reverses sign requires special treatment, which is discussed in Chapter 6.

$G^\dagger(\nu)$ applies to an outgoing wave evaluated at the LEO outside of the refracting medium. The general form for $G^\dagger(\nu)$ when $d\rho/dr > 0$ is given by

$$G^\dagger(\nu) = G[\rho^\dagger, \nu], \quad \rho^\dagger \doteq \nu - \hat{y}^\dagger K_\nu, \quad K_\nu = \left(\frac{\nu}{2}\right)^{1/3}, \quad \hat{y}^\dagger = 0.441\cdots \quad (1.5-4)$$

This connection between the outgoing and general forms for the spectral density functions is accomplished by asymptotically matching the incoming and outgoing spectral coefficients at a turning point to satisfy the general boundary condition that the spectral series shall have no singularity at the origin. It uses the property that $G[\rho, \nu] \rightarrow \text{constant}$ for decreasing $\rho < \rho^\dagger(\nu)$ to match the incoming and outgoing coefficients. The accuracy of this asymptotic matching technique, which involves an approximation, is discussed in Chapters 5 and 6.

For $\rho > \rho^\dagger(\nu)$ and provided that super-refractivity situations are avoided, it can be shown that $\partial G[\rho, \nu] / \partial \nu \doteq \tilde{\alpha}(\rho, \nu)$. Here $\tilde{\alpha}(\rho, \nu)$ is the cumulative one-way refractive bending angle evaluated at ρ for an incoming ray with an impact parameter equal to ν . Thus, $2\tilde{\alpha}(\nu, \nu) = \alpha_L(\nu)$, a virtual bending angle of a ray with an impact parameter value of ν that would be observed at the LEO. If we can recover the profile of $2G^\dagger(\nu)$ versus spectral number, then it is clear from Eqs. (1.5.3) and (1.5-4), and from the close association of $2dG^\dagger(\nu)/d\nu$ with the total virtual bending angle $\alpha_L(\nu)$ under the conditions just cited, that we can recover the profile of $d \log n / d\rho$ versus ρ , and thence $N(\rho)$. Note that scattering and reflecting surfaces that are embedded in the refracting medium, including a discontinuity in $n(r)$, can be calculated using the defining integral for $G[\rho, \nu]$ in Eq. (1.5-3). For the case of a reflecting surface, the complete spectral series breaks conveniently into two series, one for the direct field and one for the reflected field.

One technique for accomplishing the recovery of $2G^\dagger(\nu)$ is to convert the spectral series in Eq. (1.5-1) into an integral representation in phasor form using the asymptotic forms for the Hankel functions and the spherical harmonic functions. This technique also is applicable for evaluating the electromagnetic field at the LEO using numerical integration aided by the stationary-phase technique. This converted series in scalar form is given to sufficient accuracy (for $\nu < \rho_L$) by

$$\left. \begin{aligned}
 E(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho_L} \sin \theta_L} \int_0^\infty \left(\frac{\sin \theta_v^L}{\cos \theta_v^G \cos \theta_v^L} \right)^{1/2} \exp(i\Psi) dv, \\
 \Psi &= \sqrt{\rho_G^2 - v^2} + \sqrt{\rho_L^2 - v^2} + v(\theta_v^G + \theta_v^L - \theta_L) - 2G^\dagger(v) - \frac{\pi}{4}, \\
 \theta_v^G &= \sin^{-1}\left(\frac{v}{\rho_G}\right), \quad \theta_v^L = \sin^{-1}\left(\frac{v}{\rho_L}\right), \quad v < \rho_L < \rho_G
 \end{aligned} \right\} \quad (1.5-5)$$

Here $\exp(i\Psi)$ is the phasor. $\Psi = \Psi(\rho_G, \rho_L, \theta_L, v)$ gives the spectral density of the complete phase delay at the LEO position (ρ_L, θ_L) relative to the emitting GPS satellite, located at (ρ_G, θ_G) with $\theta_G = \pi$. $\Psi(\rho_G, \rho_L, \theta_L, v)$ includes the geometric delay terms and the term $-2G^\dagger(v)$ for the delay from the refractive gradient. The diagram in Fig. 1-14 shows the total geometric delay from the emitter to the receiver expressed in spectral number space. The term, $(\rho_L^2 - v^2)^{1/2}$, gives the geometric delay in phase along a straight line between the LEO and the tangency point of the line on a sphere of spectral number radius $v < \rho_L$ centered at the origin. The term $v(\theta_L - \theta_v^L)$ is an arc length along this sphere of radius v , and it is subtracted from $(\rho_L^2 - v^2)^{1/2}$ to correct it to the intersection of the sphere with the line $\theta = \pi/2$, which is the fixed reference line for computing phase delays at the LEO. Similarly, the term $(\rho_G^2 - v^2)^{1/2} + v\theta_v^G$ is the geometric phase delay from the emitting GPS satellite along the straight line to the tangency point on the sphere of radius v and thence along the sphere to the line $\theta = \pi/2$.

Although we have set the upper bound in the spectral integral to ∞ , as a practical matter the stationary-phase contributions to the integral come from

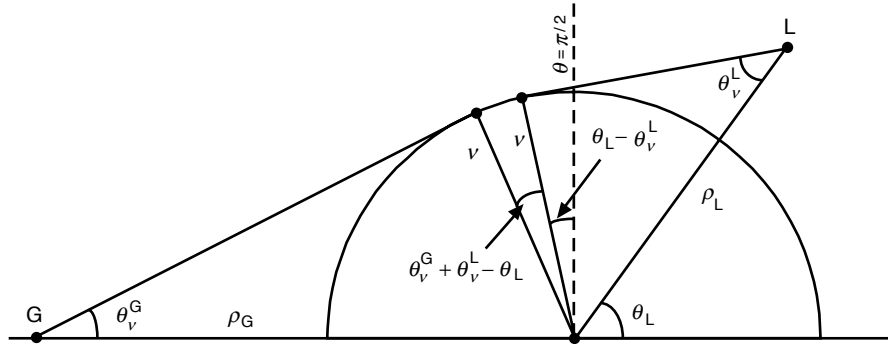


Fig. 1-14. Geometric phase delay diagram in spectral number space.

spectral numbers that are smaller than ρ_L , about 10 percent smaller for a LEO. Therefore, the asymptotic forms given in Eq. (1.5-5) for $\nu < \rho_L$ are valid.

For given satellite positions $(\rho_G, \rho_L, \theta_L)$, the stationary values of Ψ with respect to ν define the neighborhoods in spectral number that contribute to the spectral integral. When a stationary value of Ψ exists, the spectral number providing that stationary value closely equals the impact parameter (in phase units) of the corresponding ray, if near-super-refractivity conditions are avoided. In this case, the stationary value of Ψ essentially equals the value of the eikonal function for the corresponding ray, which is the phase delay φ given in Eq. (1.2-2) for the path integral along the ray (see Eq. (1.5-9) and also Appendix A for its derivation). The stationary values of Ψ with respect to spectral number involve a trade-off between geometric delay and atmospheric delay.

Next, one counter-rotates or “stops” the phase rate of $E(\rho_L, \theta_L)$ using a realistic model $\varphi_m(\rho_G, \rho_L, \theta_L)$ for the phase, which includes the Doppler shift from the LEO and GPS satellite velocities and the rate of change of the atmospheric phase delay based on geometric optics. Then the phasor in the spectral integral for $E(\rho_L, \theta_L)$ in Eq. (1.5-5) becomes $\exp[i(\Psi - \varphi_m)]$. This general phase-stopping operation is sometimes referred to as forming a hologram, the coherent mixing of two or more waves. The same holographic operation is performed on the LEO observations, and then a fast Fourier transform (or a similar discrete spectral transform) is performed. Equating these two series, the Fourier transform of the stopped LEO observations to the Fourier transform of the stopped version for $E(\rho_L, \theta_L)$ in Eq. (1.5-5), leads to an explicit evaluation of $2G^\dagger(\nu)$ versus spectral number. The resolution of this recovery is limited by the uncertainty principle in the discrete spectral analysis technique used on the data. The refractivity profile can be obtained by inverting the integral equation for $2G^\dagger(\nu)$ defined in Eqs. (1.5-3) and (1.5-4).

It is more efficient to recover the refractivity from $2dG^\dagger(\nu)/d\nu$.² It can be shown from Eq. (1.5-5) that

$$2 \frac{dG^\dagger(\nu)}{d\nu} = (\theta_v^G + \theta_v^L - \theta_{Lo}) + i\dot{\theta}_L \frac{d \log \hat{E}[\omega]}{d\omega} \bigg|_{\omega=\omega_v}, \quad \omega_v = \omega_m + \dot{\theta}_L \nu \quad (1.5-6)$$

where $d\hat{E}[\omega]/d\omega$ is the spectral derivative of the Fourier transform of the stopped $E(\rho_L, \theta_L)$, θ_{Lo} is the orbit angle of the LEO in the plane of propagation

² In wave theory, the reason for using $dG^\dagger(\nu)/d\nu$ instead of $G^\dagger(\nu)$ to recover the refractivity is similar to the reason in geometric optics for using excess Doppler instead of phase. It turns out that $\partial G[\rho, \nu]/\partial \rho = 0$ at $\rho = \rho^\dagger$, but not $\partial G[\rho, \nu]/\partial \nu$.

at the temporal center of the data interval, and $\dot{\theta}_L$ is the component of the LEO orbital angular velocity relative to the occulted GPS satellite in the plane of propagation. Note in Eq. (1.5-6) that $\omega_m \approx \pm 2 \times 10^5$ rad/s, or $\pm(30$ to $35)$ kHz; the sign depends on whether the occultation is rising or setting. On the other hand, within the time interval ΔT over which the Fourier transform is applied, $\omega_v / 2\pi$ varies over only a few tens of hertz at most. Equating $d\hat{E}[\omega_v] / d\omega_v$ in Eq. (1.5-6) to the first derivative of the Fourier transform of the LEO amplitude and stopped phase data yields a determination of the profile for $2dG^\dagger(v) / dv$. This recovered profile for $2dG^\dagger(v) / dv$ is expressed in terms of the refractivity profile through an integral equation derivable from Eqs. (1.5-3) and (1.5-4). This is given by

$$\left. \begin{aligned} 2 \frac{dG^\dagger}{dv} &\doteq 2\pi K_v \int_{\rho^\dagger}^{\infty} \left(\frac{d \log n}{d\rho} \right) \left(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \right) d\rho, \\ \hat{y} &\doteq \frac{1}{4K_v^4} (v^2 - \rho^2), \quad K_v = \left(\frac{v}{2} \right)^{1/3}, \quad \rho^\dagger = v - \hat{y}^\dagger K_v \end{aligned} \right\} \quad (1.5-7)$$

The left-hand side (LHS) of this integral equation is evaluated from the spectral derivative of the Fourier transform of the stopped LEO observations. The RHS is a convolution integral that must be inverted to obtain $\log n(\rho)$ in terms of the recovered sequence of values for $2dG^\dagger / dv$. By replacing the Airy functions with their asymptotic forms applicable for negative values of \hat{y} , $\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \rightarrow \pi^{-1}(-\hat{y})^{-1/2}$, one can show that the asymptotic version of the integral in Eq. (1.5-7) equals $\alpha_L(v)$, which is given in Eq. (1.2-6).

One can apply a wave-theory analogue of the Abel integral transform to recover $\log n(\rho)$, but in wave theory the kernel in the integral equation giving the weighting distribution of the contributions to $\log n(\rho)$ is not a Dirac delta function, as it is with the Abel transform, but rather it is spread over a characteristic width ΔL . This characteristic width corresponds to $\sim -2 \leq \hat{y} \leq 0$. Over this interval, the Airy functions make the transition from their tunneling forms to their negative argument asymptotic forms (sinusoids) corresponding to ray theory. ΔL is given by

$$\Delta L = 2 \frac{\lambda K_{\rho_o}}{2\pi} = \left(\frac{n_o r_o \lambda^2}{\pi^2} \right)^{1/3} \quad (1.5-8)$$

where r_o is the radius of curvature for the refracting surface. For GPS wavelengths at sea level, $\Delta L \approx 30$ m. Away from super-refractive areas, shadow zones, and caustics, the principal differences in results from a full-

spectrum wave theory versus a wave/optics approach originate from within this relatively narrow altitude band.

Chapter 5 establishes the correspondence between a stationary value with respect to spectral number for the spectral density function of the total delay, $\Psi^* = \Psi(\rho_G, \rho_L, \theta_L, v^*)$, when a stationary value exists, and the phase delay or eikonal function $\mathcal{S}(\rho_G, \rho_L, \theta_L, \rho_*)$ for the corresponding ray. This is the ray with an impact parameter value $\rho_* = ka = kr_*n(r_*) \doteq v^*$, where v^* is the spectral number providing a stationary value for $\Psi(\rho_G, \rho_L, \theta_L, v)$. From geometric optics, one can show [see Appendix A, Eqs. (A-55) and (A-56)] for an emitting GPS satellite located at the point (ρ_G, θ_G) with $\theta_G = \pi$, that the phase delay at the LEO position (ρ_L, θ_L) along a ray with an impact parameter value of ρ_* is given by

$$\mathcal{S} = \rho_G \cos(\chi_G + \delta_G) + \rho_L \cos(\chi_L + \delta_L) + \rho_* \alpha(\rho_*) + \int_{\rho_*}^{\infty} \alpha(\omega) d\omega \quad (1.5-9)$$

where δ_G and δ_L are the ray path deflection angles with $\alpha = \delta_G + \delta_L$, and χ_G and χ_L are orbit-related internal angles of the triangle OLG shown in Fig. 1-6. It follows from Eq. (1.5-5) that

$$\Psi(\rho_G, \rho_L, \theta_L, v^*) \Leftrightarrow \mathcal{S}(\rho_G, \rho_L, \theta_L, \rho_*), \quad v^* \doteq \rho_* \quad (1.5-10)$$

At a stationary point in spectral number, $v^* \doteq \rho_*$, and it follows from Fig. 1-14 that $\theta_{v^*}^G + \theta_{v^*}^L - \theta_L \doteq \alpha_L(\rho_*)$. But if for certain positions $(\rho_G, \rho_L, \theta_L)$ there are no stationary-phase values in spectral number for $\Psi(\rho_G, \rho_L, \theta_L, v)$, then in geometric optics there are no rays, no bending angle, and no eikonal function. This would apply to super-refractivity conditions and to strict shadow zones resulting from a discontinuity in the gradient of the refractivity (not necessarily super-refracting) or from an eclipsing limb. But even in these severe situations, $\Psi(\rho_G, \rho_L, \theta_L, v)$ still exists as a spectral density function, and it still provides a value for the field at (ρ_L, θ_L) according to Eq. (1.5-5). The field won't be zero there, just greatly diminished with damping fluctuations. Chapter 2 discusses wave-theory predictions based on scalar diffraction theory of the amplitude and phase at a LEO from various perturbations in the refractivity profile that compromise ray-theory accuracy in the transition regions (see Figs. 2-10 and 2-11). These include a strict shadow zone in amplitude caused by a discontinuity in the gradient of the refractivity, which is not super-refracting, and also by a discontinuity in refractivity, which is super-refracting.

As already mentioned, the potential radial resolution of this wave-theory technique is essentially proportional to λ / L , where L is the component of the distance traveled by the LEO in the propagation plane perpendicular to the limb

direction over ΔT , the time interval. ΔT may be limited by coherence issues; an error in ϕ_m used to stop the phase rate results in a loss of coherence in the Fourier integral transform if ΔT is too large. Horizontal variations in refractivity probably pose a larger limitation to resolution in the lower troposphere. A horizontal error δD in the assumed value of the limb distance (effectively from an unknown mesoscale horizontal refractivity gradient) translates into a vertical resolution limit that is equal to at least $500(\delta D / D)^{1/2}$ m [72,73]. A 1 percent error in D translates into a 50-m vertical resolution limit.

The formal integral for $G[\rho, \nu]$ in Eq. (1.5-3) enables one to consider a variety of features embedded in the refracting medium. In addition to a discontinuity in $n(\rho)$ or one of its derivatives, or a specular reflecting surface, or an absorbing medium by adding a complex term to $n(r)$, there are other possibilities. Spectra of recovered bending-angle profiles show broadening to varying degrees, which arise from measurement noise and atmospheric “noise” from turbulence, scintillation, and so on [29,68]. With regard to atmospheric noise, one could treat $n(\rho)$ as a stochastic variable, for example, a first-order Markoff process, with a mean value at any altitude equal to the recovered value, but with statistical parameters that are also recovered from the observations.

1.5.6 Appendices

Miscellaneous topics in geometric optics and wave theory are covered by a number of appendices. Appendix A deals principally with concepts from geometric optics and scalar diffraction theory, rays and refractive bending angle, defocusing, the Fresnel zone, the bending-angle perturbation profile from a perturbation in the refractivity profile, phase delay, and so on. Appendices A through F are intended to accompany Chapter 2. Appendix B briefly discusses a caustic surface for a ray family in terms of envelope theory. Appendix C estimates the separation altitudes of the tangency points of multipath rays as a function of the refractivity perturbation across the boundary. This includes a discussion of the effects of a discontinuity in refractivity or in one of its gradients, and the resulting caustics that can follow. Appendix D deals with third-order stationary-phase theory, an important adjunct when dealing with caustic rays in a ray-theoretic context. An index is developed for setting the accuracy of second-order geometric optics (using third-order theory) in the vicinity of a caustic contact point. Appendix E gives the bending-angle profile versus altitude for a Gaussian refractivity distribution, useful in discussing multipath, quasi-shadow zones, and caustic rays. Appendix F comments on the effect of cycle slips from either receiver operations or recovery analyses. Appendix G gives a short summary of the contour integration technique using the complex spectral number plane for summing spectral series over real integers. Appendix H develops the characteristic matrix for a stack of Airy

layers, which is discussed in Chapter 4. Appendix I summarizes the electromagnetic field equations in a stratified medium and the use of modified scalar potentials to derive the field equations. Appendix J addresses the conditions for near-equivalence between certain wave-theory phase delay quantities and their counterparts in geometric optics.

1.6 Limitations and Simplifications

This book is surely not oriented toward working with actual data. Numerous simplifying assumptions have been made to minimize the impact on an already arcane mathematical framework. Regarding the treatment of actual data from the GPS observations, regretfully we have ignored several aspects essential to use of the occultation technique for science. Not the least of our omissions is a discussion of rendering raw data streams into validated data with minimal cycle breaks, obtaining corrections to clock epoch offsets in transmitters and receivers, dealing with the attendant light-time problem to ensure that the differencing of the phase measurements to eliminate clock offsets occurs at common transmitter and/or receiver epochs, strategies for smoothing and sampling noisy data, applying estimation theory, using precision orbit determination, and, very importantly, using data information systems. The references [34,51,91,92] address many of these aspects. In addition, we have simplified the model in which GPS occultations are assumed to occur, again to minimize the impact on the mathematics. Notable examples are the following.

1.6.1 Ionosphere

Except in Chapter 2, where the thin-screen/scalar diffraction technique has been used on the sporadic E-layer, phase effects from the ionosphere have been ignored. For a comprehensive overview of ionospheric applications, see [93]. Linearly combining the dual-frequency signals from the GPS eliminates most of the ionospheric effect [93–95], but small terms involving f^{-3} and higher degrees remain that can be significant, particularly for tangency points in the upper stratosphere. Modeling the ionosphere to capture the higher-degree terms has had some success. Also, extrapolation of the dual-frequency correction from ray tangency points in the mid-troposphere to the lower troposphere (where the increased noise, particularly on the L2 phase from defocusing and interference, can result in errors in the local dual-frequency correction that exceed the correction) has been useful [59,76].

1.6.2 Placing the Occulted GPS Satellite at Infinity

Placing the emitting GPS satellite at infinity simplifies certain equations, mostly in wave theory. In geometric optics and scalar diffraction theory, this approximation can be compensated for to a certain extent by also replacing the

limb distance of the LEO, D_L , with the reduced limb distance, $D^{-1} = D_L^{-1} + D_G^{-1}$, where D_G is the actual limb distance of the occulted GPS satellite; D is about 10 percent smaller than D_L . With this simplification, the fundamental relationship between excess Doppler f_D and bending angle α of a ray is given by Eq. (1.4-5). $D\dot{\theta}_L$ is essentially the component of the orbital velocity of the LEO in the plane of propagation perpendicular to \mathbf{r}_{LG} , the direction of the GPS/LEO straight line, or, equivalently, $D\dot{\theta}_L$ is the vertical rate of descent or ascent of the GPS/LEO straight line through the atmosphere, 2 to 3 km/s. For the Earth's atmosphere, $\alpha < \sim 0.05$ rad for dry air and water vapor combined, and above the lower troposphere α is less than 0.01 rad. Thus, we have this near-linear relationship between excess Doppler and bending angle given in Eq. (1.4-5).

In contrast, for the case where the GPS satellite is at a finite distance, the appropriate form when the satellites are outside of the refracting medium is given by Eq. (1.2-3). This form can be reduced to geometric quantities in the form

$$\lambda f_D = \mathbf{T}_\perp \cdot (\mathbf{V}_L \delta_L - \mathbf{V}_G \delta_G) + O[\delta_L^2, \delta_G^2] \quad (1.6-1)$$

Here, \mathbf{T}_\perp is a unit vector in the propagation plane perpendicular to \mathbf{r}_{LG} , and δ_L and δ_G are the (small) deflection angles between the vector \mathbf{r}_{LG} and \mathbf{T}_L and \mathbf{T}_G , respectively (see Fig. 1-6). Thus, $\alpha = \delta_L + \delta_G$, and δ_L is about 8 to 9 times larger than δ_G . Nevertheless, dealing with two ray-path deflection angles instead of one (even though mutually constrained by Snell's law) and accounting for curvature effects in the approaching wave front from the occulted GPS satellite would further complicate the matter.

As described above, in wave theory a spectral integral representation is used to describe the electromagnetic wave. The main impact of adopting a finite value for \mathbf{r}_G would fall on the asymptotic form that the l th incoming spectral coefficient $a_l^-(\rho)$ would assume for large values of $\rho = knr$ out of the atmosphere. Here we have assumed that the asymptotic form for $a_l^-(\rho)$ corresponds to a collimated wave, a very simple form involving only the spectral number, $i^{l-1}(2l+1)/l(l+1)$. But with \mathbf{r}_G finite, the asymptotic form for $a_l^-(\rho)$ would correspond to a spherical wave $\exp(i\rho_{LG})/\rho_{LG}$ with its center at (r_G, θ_G) with $\theta_G = \pi$ (see Fig. 1-6). Therefore, it is an explicit function of spectral number and ρ_G through a spherical Hankel function $\xi_l^+(\rho_G)$ —not intractable, but certainly more complicated [96]. The asymptotic form of $a_l^-(\rho)$ for the case of a spherical incident wave is briefly discussed in Chapter 5, Eq. (5.5-3b). This asymptotic form involves an additional phase offset,

denoting a shift of the reference point for phase calculation from the line $\theta = \pi/2$ used in the collimated case to the position of the emitter. Also, there is an amplitude term to account for the space loss from a spherical wave. But almost all of the subsequent discussion in Chapter 5 assumes a collimated incident wave. The adjustments to account for a finite GPS satellite distance are briefly presented without derivation in Section 5.10, where ray-theory correspondence is addressed, and in Section 6.4 in connection with refractivity recovery.

1.6.3 Time

We have assumed Newtonian time throughout this monograph. But with actual data when the predicted phase must be computed for a wave moving at the speed of light, several different times and time intervals are involved: proper and coordinate times from special and general relativity, asynchronous timekeeping among all of the receiver and transmitter clocks in the system, transmit time, receive time, topocentric time, atomic time, and coordinated Universal time [34,97]. Keeping track of all these aspects of representing and synchronizing time, essential in actual data processing, unnecessarily complicates our presentation.

1.6.4 Spherical Symmetry

For the terrestrial planets with mild oblateness, the assumption of spherical symmetry is moderately accurate, but for the Jovian planets it was demonstrated about 25 years ago that oblateness must be carefully accounted for [98–100]. The usual approach for the Earth, which has a 20-km difference between its equatorial and polar radii, or 0.3 percent, is to apply the oblateness correction by adjusting the location of the geocenter using a local geoid. For a given geographical location of the tangency point of the occultation, the local radius of curvature vector for the sphere can be adjusted to fit the radius of curvature vector for the local geoid. Thus, a local spherical symmetry is assumed that more or less fits the curvature of the geopotential surface at that geographical location. The radius vector of the local sphere is equated to and aligned with the radius of curvature vector for the geopotential surface. We have assumed spherical symmetry here.

Other aspects of departures from spherical symmetry, for example, horizontal variability in the refractivity, are not considered here [38].

1.6.5 Coplanarity

The assumption of coplanarity between the propagation plane and the satellite orbit planes runs into difficulty mainly when kinematics arise. In Chapter 6, we briefly address this by identifying two different orbit angles that

locate the position of the LEO relative to the occulted GPS satellite, θ in the LEO orbit plane and $\tilde{\theta}$ in the plane of propagation. The obliquity between these two planes is readily expressed in terms of orbit elements, and it is nearly constant over an occultation episode for the neutral atmosphere. Depending on the actual orbits, the obliquity factor $d\tilde{\theta}/d\theta$ ranges between 0.6 and 1.0 for most occultations used for analysis. As discussed above, the term $d\tilde{\theta}/dt$ appears in the expressions for the excess Doppler, but it can be expressed in a near-linear form in terms of $d\theta/dt$ through spherical trigonometry when the obliquity factor is given. Also, departures from planarity from cross-track horizontal variations in refractivity are not considered here.

1.6.6 Circularity

The assumption that the LEO orbit is circular only arises here in connection with the representation of the excess Doppler in terms of the atmospheric bending and the satellite velocity. It is a fairly valid assumption, but with actual data no such approximation is made. The orbits of the LEO and the GPS satellite constellation and their clock epoch offsets are determined independently from POD information based on continuous GPS tracking data from the LEO and from a network of ground stations including the IGS.

1.6.7 Treating the GPS Signal as a Harmonic Wave

The C/A and P ranging codes on the GPS carriers are pseudorandom, phase-modulating square waves that fully suppress the carrier tone. Each transition of a code, occurring at a frequency or chip rate of 1.023 MHz for the C/A code and at 10.23 MHz for the P code, involves a change in phase of the carrier of either zero or 180 deg in accordance with the pseudorandom recipe for that particular code. Therefore, the radio frequency (RF) power spectra of the GPS navigation signals have somewhat complicated shapes. The L1 signal carries both the C/A and P codes, but on the present GPS satellites the L2 carries only the P code; however, all are phase coherent at the transmitter. Both the C/A and P codes produce $(\sin x/x)^2$ -like spread spectra with single-side bandwidths of about 1 and 10 MHz, respectively. The side lobes of the P-code spectrum at the higher harmonics of 10 MHz are attenuated at a higher rate than the $(\sin x/x)^2$ envelope decays to minimize spillover power in nearby user bands. The power in the C/A code is 3 dB greater than the power in the P code. The GPS receiver applies an appropriately time-delayed and Doppler-shifted replica of the transmitted C/A code to the received signal, and carries out a number of cross-correlation and digital signal processing tasks. Through these operations, the receiver isolates the navigation signal emitted by a specific GPS satellite from all others using the orthogonality property of the ranging codes. The receiver in essence collapses the spread-spectrum signal for each carrier

from the selected GPS satellite into a single tone, which is the equivalent of a harmonic wave with a frequency equal to that of the received carrier. Then the receiver measures the phase and amplitude of this equivalent harmonic wave in the presence of noise. Some of the operational aspects of extracting the phase and amplitude measurements from noisy digital data streams are discussed briefly in Chapter 6. For details on digital signal processing by a particular space geodetic GPS receiver, the TurboRogue, and on dealing with the encrypted P code—the so-called anti-spoofing (AS)—for recovery of the L2 phase, see [101,102]. We assume here that the received signal is indeed from harmonic waves with the L1 and L2 frequencies plus the Doppler shifts from satellite motion and from refraction. We also assume that the receiver is reporting at a suitable sample rate, nominally at 50 Hz, perhaps higher if needed, the measured phase and amplitude for this harmonic wave. Also, we assume that phase connection between successive measurements has been successful—a task not always completely achievable with actual data in adverse signal conditions.

1.7 Recommendations for the Next Chapters

Which chapters to pursue depends on one's interest. Chapter 2 and Appendices A through E have a great deal of basic material pertaining to geometric optics, scalar diffraction theory, multipath, shadow zones, caustics, stationary-phase theory, and third-order stationary-phase theory for dealing with caustics.

Chapter 3 discusses basic Mie scattering theory, but with an emphasis on a phasor-based spectral representation for the scattered wave and associated stationary-phase concepts. It can be used as a reference as needed in the later chapters for asymptotic forms, phasor representations, correspondence with ray theory concepts, and so on.

Chapter 4 stands alone. It addresses thin-film theory, which uses a unitary state transition matrix to describe electromagnetic wave propagation in a laminar medium. Several useful concepts found also in Chapters 3 and 5 are developed in Chapter 4, which offers an easier Cartesian framework for their introduction: incoming and outgoing standing waves, osculating parameters, asymptotic matching methodology, turning points, and problems therefrom. It also introduces the Airy layer wherein the refractivity gradient is constant. It is called an Airy layer because the wave equation solutions in this medium are Airy functions. It has important uses in Chapter 5 for asymptotic matching of incoming and outgoing spectral coefficients and for dealing with a turning point.

Chapter 5 provides the basis for the modified Mie scattering theory applicable to a refracting medium with or without an embedded scattering surface.

However, if one's interest is primarily in recovery issues using the full-spectrum wave-theory approach, then the main effort should be on Chapter 6 with occasional reference to Chapter 5 for the basic wave-theory fundamentals regarding the full-spectrum treatment in a spherical symmetric refracting medium.

References

- [1] C. Hazard, "The Method of Lunar Occultations and Its Application to a Survey of the Radio Source 3C212," *Monthly Notices of the Royal Astronomical Society*, vol. 124, pp. 343–357, 1962.
- [2] S. Chandrasekhar, "A Statistical Basis for the Theory of Stellar Scintillation," *Monthly Notices of the Royal Astronomical Society*, vol. 112, pp. 475–483, 1952.
- [3] W. Hubbard, J. Jokopii, and B. Wilking, "Stellar Occultations by Turbulent Planetary Atmospheres: A Wave-Optical Theory Including a Finite Scale Height," *Icarus*, vol. 34, pp. 374–395, 1978.
- [4] E. Saltpeter, "Interplanetary Scintillations. I. Theory," *The Astrophysical Journal*, vol. 147, pp. 433–448, 1967.
- [5] A. Kliore, T. Hamilton, and D. Cain, *Determination of Some Physical Properties of the Atmosphere of Mars from Changes in the Doppler Signal of a Spacecraft on an Earth Occultation Trajectory*, Technical Report 32-674, Jet Propulsion Laboratory, Pasadena, California, 1964.
- [6] G. Fjeldbo, "Bistatic Radar Methods for Studying Planetary Ionospheres and Surfaces," Science Report No. 2, NsG-377, SU-SEL-64-025, Stanford Electronics Laboratories, Stanford, California, 1964.
- [7] A. Kliore, D. Cain, G. Levy, V. Eshleman, G. Fjeldbo, and F. Drake, "Occultation Experiment: Results of the First Direct Measurement of Mars' Atmosphere," *Science*, vol. 149, pp. 1243–1248, 1965.
- [8] G. Fjeldbo, A. J. Kliore, and V. R. Eshleman, "The Neutral Atmosphere of Venus as Studied by Mariner V Radio Occultation Experiments," *The Astronomical Journal*, vol. 76, no. 2, pp. 123–140, 1971.
- [9] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120–166, 1986.
- [10] G. Lindal, J. Lyons, D. Sweetnam, V. Eshleman, D. Hinson, and G. Tyler, "The Atmosphere of Uranus: Results of Radio Occultation Measurements with Voyager 2," *Journal of Geophysical Research*, vol. 92, pp. 14,987–15,001, 1987.
- [11] G. Tyler, "Radio Propagation Experiments in the Outer Solar System with Voyager," *Proceedings of the IEEE*, vol. 75, pp. 1404–1431, 1987.

- [12] G. Lindal, "The Atmosphere of Neptune: An Analysis of Radio Occultation Data Acquired by Voyager 2," *The Astronomical Journal*, vol. 103, pp. 967–982, 1991.
- [13] O. Yakovlev, I. Matyugov, and I. A. Vilkov, "Radio-Wave Phase and Frequency Fluctuations as Observed in Radio Occultation Experiments on the Satellite-to-Satellite Link," *Journal of Communications Technology and Electronics*, vol. 41, no. 11, pp. 993–998, 1996.
- [14] A. S. Liu, *On the Determination and Investigation of the Terrestrial Ionospheric Refractive Indices Using GEOS-3/ATS-6 Satellite-to-Satellite Tracking Data; Final Report*, NASA-CR-156848, November 1, 1978.
- [15] T. Yunck, G. Lindal, and C. H. Liu, "The Role of GPS in Precise Earth Observations," *Proceedings of the IEEE: Position Location and Navigation Symposium*, Orlando, Florida, 1988.
- [16] R. Ware, M. Exner, D. Feng, M. Gorbunov, K. Hardy, B. Herman, Y.-H. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Solheim, X. Zou, R. Anthes, S. Businger, and K. Trenberth, "GPS Sounding of the Atmosphere from Low Earth Orbit: Preliminary Results," *Bulletin of the American Meteorological Society*, vol. 77, pp. 19–40, 1996.
- [17] E. Kursinski, G. Hajj, W. Bertiger, S. Leroy, T. Meehan, L. Romans, J. Schofield, D. McCleese, W. Melbourne, C. Thornton, T. Yunck, J. Eyre, and R. Nagatani, "Initial Results of Radio Occultation Observations of Earth's Atmosphere Using the Global Positioning System," *Science*, vol. 271, pp. 1107–1110, 1996.
- [18] G. A. Hajj, E. R. Kursinski, W. I. Bertiger, S. S. Leroy, T. Meehan, L. J. Romans, and J. T. Schofield, "Initial Results of GPS–LEO Occultation Measurements of Earth's Atmosphere Obtained with the GPS/MET Experiment," *Proceedings of the IAG Symposium, G1, GPS Trends in Precise Terrestrial, Airborne, and Spaceborne Applications, IUGG XXI General Assembly, Boulder, Colorado, July 2–14, 1995*, Springer-Verlag, pp. 144–154, 1996.
- [19] C. Rocken, R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.-H. Kuo, and X. Zou, "Analysis and Validation of GPS/MET Data in the Neutral Atmosphere," *Journal of Geophysical Research*, vol. 102, pp. 29,849–29,866, 1997.

- [20] J. Wickert, T. Schmidt, C. Marquardt, Ch. Reigber, K.-H. Neumayer, G. Beyerle, R. Galas, and L. Grunwaldt, "GPS Radio Occultation with CHAMP: First Results and Status of the Experiment," *Proceedings of IAG Scientific Assembly, Vistas for Geodesy in the New Millenium*, September 2–7, 2001, Budapest, Hungary, Springer Series, vol. 125, pp. 273–278, 2002.
- [21] G. Hajj, C. Ao, B. Iijima, D. Kuang, E. Kursinski, A. Mannucci, T. Meehan, L. Romans, M. de la Torre Juarez, and T. Yunck, "CHAMP and SAC-C Atmospheric Occultation Results and Intercomparisons," submitted to *Journal of Geophysical Research, Atmosphere* August 2, 2002.
- [22] R. Anthes, C. Rocken, and Y. H. Kuo, "Applications of COSMIC to Meteorology and Climate," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 115–156, 2000.
- [23] Y.-H. Kuo, X. Zou, F. Vandenberghe, W. Huang, and R. Anthes, "Assimilation of GPS/MET Data for Numerical Weather Prediction," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 157–186, 2000.
- [24] C. Rocken, Y.-H. Kuo, W. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick, "COSMIC System Description," *Terrestrial, Atmospheric and Oceanic Science*, vol. 11, pp. 21–52, 2000.
- [25] T. Yunck, C. H. Liu, and R. Ware, "A History of GPS Sounding," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 1–20, 2000.
- [26] T. P. Yunck, G. A. Hajj, E. R. Kursinski, J. A. LaBrecque, S. T. Lowe, M. M. Watkins, and C. McCormick, "AMORE: An Autonomous Constellation Concept for Atmospheric and Oceanic Observation," *Acta Astronautica*, vol. 46, no. 2-6, pp. 355–364, 2000.
- [27] B. Ahmad and G. Tyler, "The Two-Dimensional Resolution Kernel Associated with Retrieval of Ionospheric and Atmospheric Refractivity Profiles Using the Abelian Inversion of Radio Occultation Ohase Data," *Radio Science*, vol. 33, pp. 469–478, 1997.
- [28] A. Pavelyev, A. Zakharov, A. Kucherjavenkov, E. Molotov, I. Siderenko, I. Kucherjavenkova and D. Pavelyev, "Propagation of Radio Waves Reflected from Earth's Surface at Grazing Angles between a Low-Orbit Space Station and a Geostationary Satellite," *Journal of Communications Technology and Electronics*, vol. 42, no. 1, pp. 51–57, 1997.
- [29] G. Beyerle, K. Hocke, J. Wickert, T. Schmidt, C. Marquardt, and Ch. Reigber, "GPS Radio Occultations with CHAMP: A Radio Holographic Analysis of GPS Signal Propagation in the Troposphere and Surface Reflections," *Journal of Geophysical Research*, vol. 107, no. D24, 4802, doi: 10.1029/2001JD001402, 2002.

- [30] *Global Positioning System*, ed. P. M. Janiczek, vols. 1–3, Washington, DC: Institute of Navigation, 1980.
- [31] B. Parkinson and J. Spilker, eds., *Global Positioning System: Theory and Application*, American Institute of Aeronautics and Astronautics, Washington, DC, 1996.
- [32] High Speed network JPL.
- [33] R. Galas, J. Wickert, and W. Burghardt, “High Rate Low Latency GPS Ground Tracking Network for CHAMP,” *Physics and Chemistry of the Earth*, part A, vol. 26, issues 6–8, pp. 649–652, 2001.
- [34] G. Hajj, E. Kursinski, L. Romans, W. Bertiger, and S. Leroy, “A Technical Description of Atmospheric Sounding by GPS occultation,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, pp. 451–469, 2002.
- [35] J. Zumberge, and G. Gendt, “The Demise of Selective Availability and Implications for the International GPS Service,” *Physics and Chemistry of the Earth*, part A, vol. 26, issues 6–8, pp. 637–644, 2001.
- [36] G. Beutler, M. Rothacher, S. Schaer, T. A. Springer, J. Kouba, and R. E. Neilan, “The International GPS Service (IGS): An Interdisciplinary Service for Earth Sciences,” *Advances in Space Research*, vol. 23, no. 4, pp. 631–653, 1999.
- [37] R. Neilan, J. Zumberge, G. Beutler, and J. Kouba, “The International GPS Service: A Global Resource for GPS Applications and Research,” *Proc. ION-GPS-97*, The Institute of Navigation, pp. 883–889, 1997.
- [38] B. Ahmad and G. Tyler, “Systematic Errors in Atmospheric Profiles Obtained from Abelian Inversion of Radio Occultation Data: Effects of Large Scale Horizontal Gradients,” *Journal of Geophysical Research*, vol. 104, no. D4, pp. 3971–3992, 1999.
- [39] S. Sokolovskiy and D. Hunt, “Statistical Optimization Approach for GPS/MET Data Inversions,” *URSI GPS/MET Workshop*, Tucson, Arizona, February 1996.
- [40] R. Phinney and D. Anderson, “On the Radio Occultation Method for Studying Planetary Atmospheres,” *Journal of Geophysical Research*, vol. 73, no. 5, pp. 1819–1827, 1968.
- [41] G. Fjeldbo and V. Eshelman, “The Atmosphere of Mars Analyzed by Integral Inversion of the Mariner IV Occultation Data,” *Planetary Space Science*, vol. 16, pp. 1035–1059, 1968.
- [42] E. Kursinski, G. Hajj, J. Schofield, R. Linfield, and K. Hardy, “Observing Earth’s Atmosphere with Radio Occultation Measurements Using the Global Positioning System,” *Journal of Geophysical Research*, vol. 102, pp. 23,429–23,465, 1997.

- [43] O. Tetens, "Über Einige Meteorologische Begriffe," *Zeitschrift für Geophysik*, vol. 6, pp. 297–309, 1930.
- [44] J. Eyre, "Assimilation of Radio Occultation Measurements into a Numerical Weather Prediction System," ECMWF Technical Memorandum 199, European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom, 1994.
- [45] X. Zou, Y.-H. Kuo, and Y.-R. Guo, "Assimilation of Atmospheric Radio Refractivity Using a Nonhydrostatic Adjoint Model," *Monthly Weather Review*, vol. 123, pp. 2229–2249, 1995.
- [46] X. Zou, F. Vandenberghe, B. Huang, M. Gorbunov, Y.-H. Kuo, S. Sokolovskiy, J. Chang, J. Sela, and R. Anthes, "A Ray-Tracing Operator and Its Adjoint for the Use of GPS/MET Refraction Angle Measurements," *Journal of Geophysical Research, Atmospheres*, vol. 104, no. D18, pp. 22,301–22,318, 1999.
- [47] Y.-H. Kuo, X. Zou, F. Vandenberghe, W. Huang, and R. Anthes, "Assimilation of GPS/MET Data for Numerical Weather Prediction," *Terrestrial, Atmospheric and Oceanic Sciences*, 1999.
- [48] M. Born and E. Wolf, *Principles of Optics*, 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.
- [49] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, New York: McGraw-Hill, 1965.
- [50] S. Leroy, "Measurement of Geopotential Heights by GPS Radio Occultation," *Journal of Geophysical Research*, vol. 102, pp. 6971–6986, 1997.
- [51] K. Hocke, "Inversion of GPS Meteorology Data," *Annales Geophysicae*, vol. 15, pp. 443–450, 1997.
- [52] K. Hocke, T. Tsuda, and A. de la Torre, "A Study of Stratospheric GW Fluctuations and Sporadic E at Midlatitudes with Focus on Possible Orographic Effect of Andes," *Journal of Geophysical Research*, vol. 107, no. D20, 4428, doi: 10.1029/2001JD001330, October 2002.
- [53] K. Igarashi, A. Pavelyev, K. Hocke, D. Pavelyev, and J. Wickert, "Observation of Wave Structures in the Upper Atmosphere by Means of Radio Holographic Analysis of the Radio Occultation Data," *Advances in Space Research*, vol. 27, no. 6–7, pp. 1321–1326, 2001.
- [54] A. Pavelyev, K. Igarashi, Ch. Reigber, K. Hocke, J. Wickert, G. Beyerle, S. Matyugov, A. Kucherjavenkov, D. Pavelyev, and O. I. Yakovlev, "First Application of the Radio Holographic Method to Wave Observations in the Upper Atmosphere," *Radio Science*, vol. 37, no. 3, pp. 15-1–15-11, 2002.

- [55] *Atmospheric Profiling Mission*, ESA Report SP-1196, no. 7, European Space Agency, April 1996.
- [56] Y.-H. Kuo, X. Zou, and W. Huang, "The Impact of GPS Data on the Prediction of an Extratropical Cyclone: An Observing System Simulation Experiment," *Dynamics of Atmospheres and Oceans*, vol. 27, pp. 439–470, 1997.
- [57] E. K. Smith, Jr., and S. Weintraub, "The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies," *Proceedings of the IRE*, vol. 41, pp. 1035–1037, 1953.
- [58] D. Thayer, "An Improved Equation for the Radio Refractive Index in Air," *Radio Science*, vol. 9, pp. 803–807, 1974.
- [59] G. A. Hajj and L. J. Romans, "Ionosphere Electron Density Profiles Obtained with the Global Positioning System: Results from the GPS/MET Experiment," *Radio Science*, vol. 33, no. 1, pp. 175–190, 1998.
- [60] M. E. Gorbunov, S. V. Sokolovsky, and L. Bengtsson, *Space Refractive Tomography of the Atmosphere: Modeling of Direct and Inverse Problems*, Report 210, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.
- [61] T. Meehan, personal communication, Jet Propulsion Laboratory, Pasadena, California, 1995.
- [62] A. G. Pavelyev, A. V. Volkov, A. I. Zakharov, S. A. Krutikh, and A. I. Kucherjavenkov, "Bistatic Radar as a Tool for Earth Investigation Using Small Satellites," *Acta Astronautica*, vol. 39, no. 9–12, pp. 721–730, 1996.
- [63] K. Hocke, A. Pavelyev, O. Yakovlev, L. Barthes, and N. Jakowski, "Radio Occultation Data Analysis by the Radioholographic Method," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 61, pp. 1169–1177, 1999.
- [64] K. Igarashi, A. Pavelyev, K. Hocke, D. Pavelyev, I. Kucherjavenkov, S. Matyugov, A. Zakharov, and O. Yakovlev, "Radio Holographic Principle for Observing Natural Processes in the Atmosphere and Retrieving Meteorological Parameters from Radio Occultation Data," *Earth, Planets and Space*, vol. 52, no. 11, pp. 893–899, 2000.
- [65] M. E. Gorbunov, A. S. Gurvich, and L. Kornbluh, "Comparative Analysis of Radioholographic Methods of Processing Radio Occultation Data," *Radio Science*, vol. 35, no. 4, pp. 1025–1034, 2000.
- [66] S. Sokolovsky, "Modeling and Inverting Radio Occultation Signals in the Moist Troposphere," *Radio Science*, vol. 36, no. 3, pp. 441–458, 2001.

- [67] A. Pavelyev, Y.-A. Liou, Ch. Reigber, J. Wickert, K. Igarashi, K. Hocke, and C.-Y. Huang, "GPS Radio Holography as a Tool for Remote Sensing of the Atmosphere and Mesosphere from Space," *GPS Solutions*, vol. 6, pp. 100–108, 2002.
- [68] K. Igarashi, K. Pavelyev, J. Wickert, K. Hocke, and D. Pavelyev, "Application of Radio Holographic Method for Observation of Altitude Variations of the Electron Density in the Mesosphere/Lower Thermosphere Using GPS/MET Radio Occultation Data," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, no. 8–11, pp. 959–969, 2002.
- [69] O. I. Yakovlev, A. G. Pavelyev, S. S. Matyugov, D. A. Pavelyev, and V. A. Anufriev, *Radioholographic Analysis of Radiooccultation Data*, Scientific Technical Report 02/05, GeoForschungsZentrum-Potsdam, Potsdam, Germany, 2002.
- [70] J. Goodman, *Introduction to Fourier Optics*, New York: McGraw-Hill, 1968.
- [71] E. N. Bramley, "The Accuracy of Computing Ionospheric Radiowave Scintillation by the Thin-Phase Screen Approximation," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 39, pp. 367–373, 1977.
- [72] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120–166, 1986.
- [73] W. G. Melbourne, E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, E. R. Kursinski, T. K. Meehan, L. E. Young, and T. P. Yunck, *The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring*, JPL Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, April 1994.
- [74] M. Meincke, "Inversion Methods for Atmospheric Profiling with GPS Occultation," Scientific Report No. 99-11, Danish Meteorological Institute, Copenhagen, Denmark, 1999.
- [75] A. J. Devaney, "A Filtered Backpropagation Algorithm for Diffraction Tomography," *Ultrasonic Imaging*, vol. 4, pp. 336–350, 1982.
- [76] M. E. Gorbunov, A. S. Gurvich, and L. Bengtsson, *Advanced Algorithms of Inversion of GPS/MET Satellite Data and Their Application to Reconstruction of Temperature and Humidity*, Report 211, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.
- [77] E. Karayel and D. Hinson, "Sub-Fresnel Scale Vertical Resolution in Atmospheric Profiles from Radio Occultation," *Radio Science*, vol. 32, no. 2, pp. 411–423, 1997.

- [78] R. H. Hardin and F. D. Tappert, “Applications of the Split-Step Fourier Method to the Numerical Solution of Nonlinear and Variable Coefficient Wave Equations,” *SIAM Review*, vol. 15, p. 423, 1973.
- [79] D. Knepp, “Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves,” *Proceedings of the IEEE*, vol. 71, no. 6, pp. 722–737, 1983.
- [80] J. Martin and S. Flatté, “Intensity Images and Statistics from Numerical Simulation of Wave Propagation in 3-D Random Media,” *Applied Optics*, vol. 27, no. 11, pp. 2111–2126, 1988.
- [81] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, United Kingdom: The Institution for Electrical Engineers, 2000.
- [82] G. Beyerle, M. Gorbunov, and C. Ao, “Simulation Studies of GPS Radio Occultation Measurements,” *Radio Science*, in press, 2003.
- [83] C. Ao, G. Hajj, T. Meehan, S. Leroy, E. Kursinski, M. de la Torre Juárez, B. Iijima, and A. Mannucci, “Backpropagation Processing of GPS Radio Occultation Data,” in *CHAMP Science Meeting*, Ch. Reigber, H. Lühr, and P. Schwintzer, eds., January 21–24, 2002, Springer-Verlag, Berlin, Germany, pp. 415–422, 2003.
- [84] M. E. Gorbunov, “Canonical Transform Method for Processing Radio Occultation Data in the Lower Troposphere,” *Radio Science*, vol. 37, no. 5, 1076, doi:10.1029/2000RS002592, 2002.
- [85] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Series 55, Washington, DC, 1964.
- [86] F. Olver, *Introduction to Asymptotics and Special Functions*, New York: Academic Press, 1974.
- [87] W. G. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occultation Observations from a Low Earth Orbiter, Part 1*, JPL Publication 99-5, Jet Propulsion Laboratory, Pasadena, California, July 1998.
- [88] W. G. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occultation Observations from a Low Earth Orbiter, Part 2*, JPL Publication 99-5, Jet Propulsion Laboratory, Pasadena, California, April 1999.
- [89] F. Abelès, “Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Applications to Thin Films,” *Annales de Physique*, vol. 5, pp. 596–640, September–October 1950 (in French).

- [90] F. Abelès, “Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Application to Thin Films. II. Thin Films,” *Annales de Physique*, vol. 5, pp. 706–782, November–December 1950 (in French).
- [91] J. Wickert, G. Beyerle, T. Schmidt, C. Marquardt, R. König, L. Grunwald, and C. Reigber, “GPS Radio Occultation with CHAMP,” *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*, eds. C. Reigber, H. Luhr, and P. Schwintzer, proceedings from the First CHAMP Science Meeting, Potsdam, Germany, January 21–24, 2002, Springer-Verlag, 2002.
- [92] <http://genesis.jpl.nasa.gov/>
- [93] G. Hajj, L. Lee, X. Pi, L. Romans, W. Schreiner, P. Strauss, and C. Wang, “COSMIC GPS Ionosphere Sensing and Space Weather,” *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 235–272, 2000.
- [94] S. Bassiri and G. Hajj, “Higher-Order Ionospheric Effects on the Global Positioning System Observables and Means of Modeling Them,” *Manuscripta Geodaetica*, vol. 18, pp. 280–289, 1993.
- [95] V. V. Vorob’ev and T. G. Krasil’nikova, “Estimation of the Accuracy of the Atmospheric Refractive Index Recovery from Doppler Shift Measurements at Frequencies Used in the NAVSTAR System,” *Physics of the Atmosphere and Ocean*, vol. 29, no. 5, pp. 602–609, 1994.
- [96] J. Jackson, *Classical Electrodynamics*, 2nd ed., New York: John Wiley & Sons, Inc., 1975.
- [97] T. D. Moyer, *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation*, Monograph 2, Deep Space Communications and Navigation Series, Jet Propulsion Laboratory, Pasadena, California, October 2000.
- [98] W. Hubbard, “Ray Propagation in Oblate Atmospheres,” *Icarus*, vol. 27, pp. 387–388, 1976.
- [99] A. J. Kliore, P. M. Woiceshyn, and W. B. Hubbard, “Temperature of the Atmosphere of Jupiter from Pioneer 10/11 Radio Occultations,” *Geophysical Research Letters*, vol. 3, pp. 113–116, 1976.
- [100] G. F. Lindal, D. N. Sweetnam, and V. R. Eshleman, “The Atmosphere of Saturn: An Analysis of the Voyager Radio Occultation Measurements,” *The Astronomical Journal*, vol. 90, pp. 1136–1146, 1985.
- [101] J. B. Thomas, *Signal-Processing Theory for the TurboRogue Receiver*, JPL Publication 95-6, Jet Propulsion Laboratory, Pasadena, California, 1995.

- [102] C. Ao, T. Meehan, G. Hajj, A. Mannucci, and G. Beyerle, “Lower Troposphere Refractivity Bias in GPS Occultation Retrievals,” submitted to *Journal of Geophysical Research*, 2003.
- [103] T. Meehan (operational aspects of TurboRogue).

Chapter 5

Propagation and Scattering in a Spherical-Stratified Refracting Medium

5.1 Introduction

We extend the Mie scattering treatment in Chapter 3 to the case of a scattering spherical surface embedded in a refracting medium that is laminar but otherwise spherical symmetric. Figure 5-1 depicts this scenario. At the boundary, n or one of its derivatives may be discontinuous with a refraction profile $n^+(r)$ in the overlying medium and $n^-(r)$ inside the sphere. The angle α_L is the refractive bending angle at the LEO. The angle $\tilde{\alpha}$ is the cumulative bending angle along the ray. A thin phase screen approach to this problem is discussed in Chapter 2. There the scalar diffraction integral, which is a convolution integral over the vertical continuum of emitters in a thin screen model, is used to calculate the phase and amplitude of the emitted wave at a point some distance from the thin screen. In that essentially wave/optics approach, the thin screen serves as a surrogate for the actual atmosphere. The radiation field from the emitters in the screen mimics the phase and amplitude effects on the electromagnetic wave resulting from propagating through the actual atmosphere, including the effects from its refractive gradient and from the embedded scattering surface. Fresnel diffraction, interference, shadowing, caustics, etc., all can be evaluated using the thin phase screen approach combined with the scalar diffraction integral.

A wave theory approach is based on solutions to Maxwell's equations applied to a spherical atmosphere. A convolution integral also appears, but it is over spectral number instead of thin screen altitude. The wave theory approach can be considered as more accurate, albeit computationally more expensive. In both approaches, thin screen/scalar diffraction theory and wave theory, one

ends up with a prediction of the observed phase and amplitude of the wave at some point as a result of its passing through an intervening atmosphere and perhaps encountering an embedded scattering surface. One question addressed here is the level of agreement between these two approaches, and how that level depends on the adversity of the wave propagation conditions in the atmosphere.

One concludes from a review of wave propagation literature that scattering theory in a sphere is potentially a very complicated problem. For example, see the survey article by Chapman and Orcutt [1] on wave propagation problems in seismology. There one finds refracted rays reflecting from multiple surfaces, Rayleigh and Love waves skittering along boundaries, super-refracted waves with multiple reflections ducted along between layers, and so on. Here we specifically rule out ducting, evanescence, or other confounding propagation effects, except for the effects resulting from the class of discontinuities under study here, which would include, however, interference, shadow zones, caustics, diffraction and super-refractivity. We assume that embedded in and co-centered with this refracting medium is a single large spherical scattering surface. Across this surface a discontinuity in the refractivity model is assumed to exist. This discontinuity can take different forms ranging from a discontinuity in n itself or in its gradient, or merely in one of its higher derivatives. We must account for the effects of the refractive gradient in the overlying medium surrounding the sphere on the phase and amplitude of the electromagnetic wave. Therefore, scattering in this context includes external reflection from the scattering sphere, transmission through and refraction by the scattering sphere, including the possibility that the scattered wave has undergone one or more reflections inside the scattering sphere, and finally the refractive bending of the scattered ray from the overlying medium.

For the purpose of comparing results from the full wave theory approach with the scalar diffraction/thin phase screen approach, we assume that the local gradient of the refractivity is sufficiently small throughout the medium surrounding the scattering sphere so that the “thin atmosphere” conditions (see Section 2.2, Eqs. (2.2-8) and (2.2-9)) apply. Where rapid changes in refractivity are encountered, for example, at the boundary of a super-refracting water vapor layer, we assume that such changes are sufficiently localized so that ray optics is still valid, i.e., rays do exist that pass through such a barrier, for at least a certain range of tangency points.

The wave theory approach followed here is derived from Mie scattering theory, but it is adapted to a medium with a continuously changing refractivity. The original formulation of Mie scattering theory [2] deals with a single spherical scattering surface in an otherwise homogeneous medium. Numerical wave theory approaches involve approximations to the solutions of Maxwell’s equations in one form or another. In this chapter we use an osculating parameter technique for dealing with the spectral integrals associated with wave theory. The accuracy of such a technique and its range of applicability are

important questions that need to be addressed. The accuracy and range of applicability depend on the choice of basis functions used in the osculating parameter technique. For example, in a Cartesian-stratified medium, the use of sinusoids as the basis functions results in an osculating parameter solution that is identical to the WKB solution. The conditions for attaining a given accuracy and ascertaining its range of applicability are well established for WKB solutions. There also is a wealth of literature on the connection problem in WKB solutions across the transition zone between the oscillatory and exponential-decaying branches, important for quantum tunneling processes and other applications. Fortunately, we are concerned with the electromagnetic field away from turning points; therefore, asymptotic forms applicable to the oscillatory branch play an important role here. For a different choice of basis functions the osculating parameter solutions do not reduce to the WKB forms and have a different range of applicability. Here the favored basis functions are the spherical Bessel functions, or their Airy function surrogates, which are asymptotically equivalent when the radius of the scattering surface is very large compared to the wavelength of the electromagnetic wave. These particular basis functions offer a wide range of applicability for the osculating parameter solutions. Even at a turning point, a *bête noir* for wave theory, these basis functions provide a useful, if not completely successful approach.

The question arises concerning the many sections to follow as to which parts are essential to this wave theory approach. Sections 5.2 and 5.3 provide a brief review of the basic general concepts in classical electrodynamics involving harmonic waves: Maxwell's equations, scalar potentials for generating the electromagnetic field vectors, series solutions using the separation of variables technique. These series involve spherical harmonic functions, which apply for spherical symmetry, but special functions are needed for the radial component. For the homogeneous case these radial solutions become the spherical Bessel functions, but in general the radial functions depend on the refractivity profile of the medium. It is here that techniques like the WKB method or the osculating parameter technique arise.

Section 5.4 briefly summarizes the asymptotic approximations that are used in this chapter. A fuller account is found in Section 3.8. This section is referred to frequently in the later text.

Section 5.5 begins the adaptation of Mie scattering theory from a single large spherical surface to a concatenated series of concentric layers that in its limiting form approach a medium with a continuously varying refractivity. This section introduces a spectral density function for the phase delay induced by the refractive gradient in the medium. This quantity (defined as $G[\rho, \nu]$ in that section) essentially accounts for the extra phase delay at the radial position ρ experienced by a radial wave component of spectral number $l = \nu - 1/2$, which results from the refractive gradient of the medium. In a homogeneous medium

$G[\rho, \nu] \equiv 0$. We consider there the propagation of an incident wave that asymptotically is planar at large approaching distances relative to the scattering sphere. The adjustments to account for a spherical approaching wave (when the emitting GPS satellite is placed only a finite distance away) are noted.

Section 5.6 reviews several important concepts from geometric optics that are needed later where correspondences are established between these concepts and certain properties from wave theory when the spectral number assumes a stationary phase value. Geometric optics is discussed in Appendix A, but here, in addition to discussing the stationary phase property of a ray and its bending angle and phase delay, this section also introduces the concept of a cumulative bending angle along the ray, which mainly arises when evaluating the electromagnetic field within the refractive medium. This section also discusses Bouguer's law and the impact parameter of a ray, the geometric optics equivalent of the conservation of angular momentum in a conservative force field. This section also covers defocusing and the first Fresnel zone. Limitations in second order geometric optics, which arise in association with caustics or when two or more rays have impact parameter separations that are less than the first Fresnel zone, are discussed in Section 5.12.

Section 5.7 develops more asymptotic forms needed in the sequel. The ratio of the radius of curvature of the stratified surface to the wavelength of the incident wave, r_o / λ , is sufficiently large so that asymptotic forms for the Bessel functions apply. Also, because only spectral numbers that are of the same order of magnitude as r_o / λ contribute significantly to the spectral integrals representing the field, we also can use the asymptotic forms for the harmonic functions that apply for large spectral number. Section 5.7 shows the close correspondence between certain geometric optics quantities, for example, the cumulative bending angle of a ray with an impact parameter value ν and evaluated at a radial position ρ , and a certain spectral quantity from wave theory, $\partial G[\rho, \nu] / \partial \nu$. The issue of the breakdown in accuracy of the osculating parameter technique near a turning point $\nu = \rho$ also is addressed here. Guidance from the behavior of the WKB solutions near a turning point is used to deal with this breakdown. An asymptotic matching technique is developed to set the value of $G[\rho, \nu]$ for the regime $\nu > \rho$.

Section 5.8 begins the representation of the electromagnetic field in terms of the spectral integrals involving the spectral components of the radial osculating parameter functions and the harmonic functions for the angle coordinates.

This discussion is continued in Section 5.9 where a phasor representation for the integrands in these spectral integrals is introduced. The stationary phase technique also is introduced here. It is used to determine spectral number points that yield stationary values of the phasor, thereby aiding the numerical evaluation of the spectral integrals.

Section 5.10 compares results from wave theory with results from a thin phase screen model combined with the scalar diffraction integral. Correspondences between stationary phase values of certain wave theoretic quantities and their analogs in geometric optics are discussed.

Section 5.11 deals the turning point problem using an Airy layer.

Section 5.12 discusses caustics and multipath from a wave theoretic point of view in a spectral number framework. It also discusses caustics and multipath in a second order geometric optics framework, including its shortcomings near caustics or in dealing with ray pairs with nearly merged impact parameters. Third order stationary phase theory is introduced to develop a ray theory that can accurately deal with these situations. Beginning in Section 5.12 numerical solutions for the spectral representation of the field at the LEO are presented. Here the numerical integrations have been aided by the stationary phase technique to identify contributing neighborhoods in spectral number, greatly improving its efficiency.

Section 5.13 deals with a spherical scattering surface embedded in an overlying refracting medium.

Finally, Section 5.14 discusses the perfectly reflecting sphere that is embedded in an overlying refracting medium. This section also discusses the correspondence between geometric optics quantities and wave theory quantities when stationary phase values are used in each system. For example, the stationary phase values in spectral number in wave theory correspond to ray path impact parameter values that satisfy the law of reflection.

5.2 Maxwell's Equations in a Stratified Linear Medium

We follow closely the development given in Section 3.2 for the homogeneous case; the relevant symbols are defined in that section (See also the Glossary). Here Gaussian units are used. A harmonic electromagnetic wave may be written in the form

$$\mathbf{E} = \mathbf{E}(\mathbf{r})\exp(-i\omega t), \quad \mathbf{H} = \mathbf{H}(\mathbf{r})\exp(-i\omega t) \quad (5.2-1)$$

Maxwell's equations for the time-independent components in a linear medium, free of charge and current densities, are given by

$$\left. \begin{array}{ll} \text{a) } \nabla \times \mathbf{E} = ik\mu\mathbf{H}, & \text{b) } \nabla \times \mathbf{H} = -ik\varepsilon\mathbf{E}, \\ \text{c) } \nabla \cdot (\varepsilon\mathbf{E}) = 0, & \text{d) } \nabla \cdot (\mu\mathbf{H}) = 0 \end{array} \right\} \quad (5.2-2)$$

Here ε is the electrical permittivity of the propagation medium, μ is its magnetic permeability, and $k = 2\pi / \lambda$. k is the wavenumber of the harmonic wave in a vacuum, $\omega = kc$, where c is the velocity of light. These equations in Eq. (5.2-2) may be recast through successive vector calculus operations into

separate vector wave equations that \mathbf{E} and \mathbf{H} must individually satisfy [3]. These are given by

$$\left. \begin{aligned} \text{a) } \nabla^2 \mathbf{E} + \mu \epsilon k^2 \mathbf{E} + \nabla(\log \mu) \times (\nabla \times \mathbf{E}) + \nabla[\mathbf{E} \cdot (\nabla(\log \epsilon v))] &= 0 \\ \text{b) } \nabla^2 \mathbf{H} + \mu \epsilon k^2 \mathbf{H} + \nabla(\log \epsilon) \times (\nabla \times \mathbf{H}) + \nabla[\mathbf{H} \cdot (\nabla(\log \mu v))] &= 0 \end{aligned} \right\} \quad (5.2-3)$$

Here the identity $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ is used. These are the modified wave equations that the time-invariant component of a harmonic wave must satisfy in a linear medium.

We assume now that the medium is spherical stratified. In this case the index of refraction is a function of only the radial coordinate,

$$n(r) = \sqrt{\mu(r)\epsilon(r)} \quad (5.2-4)$$

It follows for this case that the gradient vectors of $\epsilon(r)$ and $\mu(r)$ are radial directed, which simplifies Eq. (5.2-3).

For the special case where $\mu \equiv 1$ throughout the medium, which is essentially the case for L-band radio signals in the neutral atmosphere, Eq. (5.2-3) is further simplified. In the special case where \mathbf{E} is perpendicular or transverse to \mathbf{r} , which is the so-called TE wave, then $\nabla \epsilon \cdot \mathbf{E}_{\text{TE}} \equiv 0$ and Eq. (5.2-3a) becomes

$$\nabla^2 \mathbf{E}_{\text{TE}} + n^2 k^2 \mathbf{E}_{\text{TE}} = 0 \quad (5.2-5)$$

Eq. (5.2-5) is nearly the Helmholtz equation (see Section 3.2, Eq. (3.2-1c)) except for the radial dependency of $n(r)$. This variation of $n(r)$ will be very slight in our case of a thin atmosphere, except possibly at a boundary. But, because r_o / λ is so large, even a small variation δn results in a significant change $k \delta n$ in the gradient of the phase accumulation of the wave.

5.2.1 Scalar Potential Functions

Following the approach in Section 3.2 for Mie scattering theory, we use the scalar potential functions for the electromagnetic field in a stratified medium expressed as a series summed over integer spectral number. It is convenient to express the electromagnetic field vectors in terms of vector calculus operations on a pair of scalar potentials, ${}^e\Pi(r, \theta, \phi)$ and ${}^m\Pi(r, \theta, \phi)$. In Section 3.2 it is shown [3] for the case where ϵ and μ are constant that these two scalar potentials are linearly independent solutions to the Helmholtz equation

$$\nabla^2 \Pi + k^2 n^2 \Pi = 0 \quad (5.2-6)$$

where, in Section 3.2, n is a constant. Each solution for a homogeneous medium can be represented using the technique of separation of variables in spherical coordinates as a series expansion. The series is expressed in terms of spherical Bessel functions of integer order l , which are a function of the radial coordinate $\rho = nkr$, and the spherical harmonic functions of degree m and order l , which are functions of the angular coordinates θ and ϕ . Here θ is the angle between \mathbf{r} and the z -axis. The latter is the axis of propagation (the asymptotic direction of the Poynting vector \mathbf{S} (Figure 5-1)) for the approaching wave. Also, ϕ is the azimuthal angle about the z -axis. See Figure 4-10 for the definition of the coordinate frame.

The electromagnetic field vectors for the homogeneous medium are obtained from a particular vector form for the scalar potentials due to Hertz. These are given by

$$\begin{aligned} \mathbf{E} &= \nabla \times \nabla \times ({}^e\Pi\mathbf{r}) + ik\mu\nabla \times ({}^m\Pi\mathbf{r}) \\ \mathbf{H} &= \nabla \times \nabla \times ({}^m\Pi\mathbf{r}) - ik\varepsilon\nabla \times ({}^e\Pi\mathbf{r}) \end{aligned} \quad (5.2-7)$$

In the electrodynamics literature, the “TE” and “TM” waves are generated from linearly independent solutions to the Helmholtz equation in Eq. (5.2-6). In Eq. (5.2-7) the term $ik\mu\nabla \times [{}^m\Pi\mathbf{r}]$ generates the electric field \mathbf{E}_{TE} , which is perpendicular to \mathbf{r} , that is, a transverse electric field; this wave is known in the literature as the “TE” wave. Similarly, the term $-ik\varepsilon\nabla \times [{}^e\Pi\mathbf{r}]$ generates a transverse magnetic field \mathbf{H}_{TM} or the “TM” wave. One can readily show (see Appendix I) that these expressions in Eq. (5.2-7) yield field vectors that satisfy Maxwell’s equations in Eq. (5.2-2) when ε and μ are constant.

For the stratified medium with $n = n(r)$, the scalar potentials are solutions to a modified Helmholtz equation. In classical electrodynamics there is a certain degree of arbitrariness in the definition of the scalar electric potential Φ and the vector magnetic potential \mathbf{A} from which \mathbf{E} and \mathbf{H} are derived. Specifically, the electromagnetic field remains invariant if Φ and \mathbf{A} are transformed together to some other pair of functions through a so-called gauge transformation, that is, the transformation is effected while Φ and \mathbf{A} are constrained to satisfy a gauge condition such as that provided by the Lorentz condition [4]. The electromagnetic field is called gauge invariant. It is rooted in the symmetries in the electrodynamics equations when they are expressed in the space-time framework of Special Relativity. There, the form of the electrodynamics equations for the 4-vector (Φ, \mathbf{A}) remains invariant under a Lorentz transformation; the 4-vector (Φ, \mathbf{A}) is called covariant in a relativistic framework.

Similarly, the scalar potentials for the stratified medium have some degree of freedom in their definition. For the case where $n = n(r)$ it is shown in

Appendix I that the electromagnetic field can be expressed through vector calculus operations on the modified scalar potentials, $[\epsilon \Pi \epsilon^{1/2} \mathbf{r}]$ and $[\mu \Pi \mu^{1/2} \mathbf{r}]$. These expressions are given by

$$\begin{aligned} \mathbf{E} &= \epsilon^{-1} \nabla \times \nabla \times [\epsilon \Pi \epsilon^{1/2} \mathbf{r}] + ik \nabla \times [\mu \Pi \mu^{1/2} \mathbf{r}] \\ \mathbf{H} &= \mu^{-1} \nabla \times \nabla \times [\mu \Pi \mu^{1/2} \mathbf{r}] - ik \nabla \times [\epsilon \Pi \epsilon^{1/2} \mathbf{r}] \end{aligned} \quad (5.2-8)$$

The factors $\epsilon^{1/2}$ and $\mu^{1/2}$ have been inserted into the potential terms in Eq. (5.2-8) to simplify the resulting modified Helmholtz equation that each of the scalar potentials must satisfy. These scalar potentials must satisfy modified Helmholtz equations, which are given by

$$\nabla^2 \epsilon \Pi + k^2 \tilde{n}_{\text{TM}}^2 \epsilon \Pi = 0, \quad \nabla^2 \mu \Pi + k^2 \tilde{n}_{\text{TE}}^2 \mu \Pi = 0 \quad (5.2-9)$$

Here the modified indices of refraction are

$$\begin{aligned} \tilde{n}_{\text{TM}}^2 &= n^2 \left[1 - \frac{r^2 \epsilon^{1/2}}{k^2 n^2} \frac{d}{dr} \left(\frac{1}{r^2} \frac{d}{dr} \left(\frac{1}{\epsilon^{1/2}} \right) \right) \right] \\ \tilde{n}_{\text{TE}}^2 &= n^2 \left[1 - \frac{r^2 \mu^{1/2}}{k^2 n^2} \frac{d}{dr} \left(\frac{1}{r^2} \frac{d}{dr} \left(\frac{1}{\mu^{1/2}} \right) \right) \right] \end{aligned} \quad (5.2-10)$$

For the case where $\mu(r) \equiv 1$ throughout the medium, we note from Eq. (5.2-10) that the modified index of refraction for the *TE* wave reduces to the regular index of refraction. When the conditions $|\nabla n| \ll 1$ and $kr_0 \gg 1$ apply, which do apply for L-band propagation in the Earth's thin neutral atmosphere, it follows that $\tilde{n}(r)$ differs from $n(r)$ by a small amount of the order of n''/k^2 ; for L-band signals in dry air this translates into a fractional difference in refractivity of roughly parts in 10^{11} . So, for computations in neutral atmosphere conditions we may simply use $n(r)$ in the modified Helmholtz wave equation. Therefore, we herewith drop the distinction between \tilde{n}_{TE} or \tilde{n}_{TM} and $n(r)$, and simply use $n(r)$ in the modified Helmholtz equation in the following discussion. For the ionosphere these differences may be more significant.

5.3 Modified Spherical Bessel Functions

We assume now that our stratified medium satisfies the asymptotic condition $n(r) \rightarrow 1$ as $r \rightarrow \infty$, so that the scalar potential series solutions for the homogeneous medium in Section 3.2 can be used as asymptotic boundary

conditions for the stratified case. For the stratified medium we again seek solutions to the modified Helmholtz equation using the technique of separation of variables of the form

$$\Pi = R(r)\Theta(\theta)\Phi(\phi) \quad (5.3-1)$$

where $\Pi(r, \theta, \phi)$ may be taken as any spectral component of a scalar potential. For large values of r , where the $n(r) \rightarrow 1$, we know that these solutions must approach the forms given in Chapter 3 for the homogeneous case. We also conclude because of the spherical symmetry of the propagation medium that spherical harmonic functions will be applicable, that is, the $\Theta(\theta)$ functions will be the same associated Legendre polynomials $P_l^m(\cos\theta)$ of order l and degree m , and the $\Phi(\phi)$ functions will be sinusoids of the form $\exp(\pm im\phi)$. For an electromagnetic vector field, the m values are restricted to $m = \pm 1$. This follows from Bauer's identity, applicable to a plane wave in a homogeneous medium (see, for example, Chapter 3, Eq. (3.2-3)), also [3]. Referring to Figure A-3, Bauer's identity is obtained from the multipole expansion [4] for a spherical wave centered at the point G and evaluated at the point L. The amplitude and phase of the time-independent part of the spherical wave is given by $\exp(inkr_{LG})/(nkr_{LG})$. Its expansion in terms of spherical Bessel and spherical harmonic functions of the transmitter and receiver coordinates is given by

$$\left. \begin{aligned} \frac{\exp(i\rho_{LG})}{\rho_{LG}} &= i \sum_{l=0}^{\infty} (2l+1) \frac{\psi_l(\rho_L)}{\rho_L} \frac{\xi_l^+(\rho_G)}{\rho_G} P_l(\cos(\theta_G - \theta_L)), \\ \rho_L < \rho_G, \quad \rho &= nkr \end{aligned} \right\} \quad (5.3-2)$$

which is obtained by applying the addition theorem for spherical harmonic functions. If we now let $\rho_G \rightarrow \infty$ and $\theta_G = \pi$, then $\rho_{LG} \rightarrow \rho_G + \rho_L \cos\theta_L$ and we can replace $\xi_l^+(\rho_G)$ with its asymptotic form for large $\rho_G \gg l$, $\xi_l^+(\rho_G) \rightarrow (-i)^{l+1} \exp(i\rho_G)$. We substitute these forms into the above expansion for the spherical wave, cancel terms and note that $P_l(-x) = (-1)^l P_l(x)$. It follows that for a plane harmonic wave traveling along the z -axis in a homogeneous medium, the time-independent component is given by $\exp(i\rho \cos\theta)$, and that Bauer's identity is given by

$$\exp(i\rho \cos\theta) = \sum_{l=0}^{\infty} i^l (2l+1) \frac{\psi_l(\rho)}{\rho} P_l(\cos\theta), \quad \rho = nkr \quad (5.3-3)$$

The vector version is given by multiplying Eq. (5.3-3) by $(\hat{r} \sin \theta \cos \phi + \hat{\theta} \cos \theta \cos \phi + \hat{\phi} \sin \theta)$. When the coefficients of the basis functions $R(r, l) P_l^m(\cos \theta) \exp(\pm i m \phi)$ in the series solution to Eq. (5.2-6) for a given vector component of the field are matched on a term-by-term basis with the corresponding coefficients in the Bauer series for the same vector component (and using the property $\partial P_l / \partial \theta = -P_l^1$), one finds that m is indeed restricted to $m = \pm 1$. This restriction is perpetuated to the scattered field by the continuity conditions in electrodynamics that apply to the field components across a scattering boundary¹. The form of the series solutions in this case must approach as $r \rightarrow \infty$ the same form given in Chapter 3, Eq. (3.2-4). Only the $R(r)$ functions will differ from the spherical Bessel functions that apply to the homogeneous case, and these modified functions will approach the Bessel function form as $r \rightarrow \infty$. Thus, we have

$$rR = \tilde{\xi}_l^\pm(kr) \quad (5.3-4)$$

where $\tilde{\xi}_l^\pm$ is related to the spherical Hankel functions of the first (+) and second (-) kind, but modified for the stratified medium. These functions must satisfy the modified differential equation for spherical Bessel functions, which is given by

$$\tilde{\xi}_l^{\pm''} + \left(n(r)^2 - \frac{l(l+1)}{u^2} \right) \tilde{\xi}_l^\pm = 0 \quad (5.3-5)$$

Here, $u = kr$ and $(*)' = d(*)/du$. See Chapter 3, Section 3.2, Eq. (3.2-8) for the definition of these spherical Hankel functions in the case of a homogenous medium in terms of the integer Bessel functions of the first and second kind. In particular, the relationships between the modified spherical Bessel functions of the first and second kind, $\tilde{\psi}_l(u)$ and $\tilde{\chi}_l(u)$, for the stratified medium and the modified spherical Hankel functions are given by

$$\left. \begin{aligned} \tilde{\xi}_l^+(u) &= \tilde{\psi}_l(u) + i\tilde{\chi}_l(u), & \tilde{\xi}_l^-(u) &= \tilde{\psi}_l(u) - i\tilde{\chi}_l(u), \\ \tilde{\psi}_l(u) &= \frac{1}{2}\tilde{\xi}_l^+(u) + \frac{1}{2}\tilde{\xi}_l^-(u), & \tilde{\chi}_l(u) &= \frac{1}{2i}\tilde{\xi}_l^+(u) - \frac{1}{2i}\tilde{\xi}_l^-(u) \end{aligned} \right\} \quad (5.3-6)$$

For $u \rightarrow 0$, $\tilde{\psi}_l(u) \rightarrow 0$ and $\tilde{\chi}_l(u) \rightarrow \infty$.

¹ The quantum mechanical analog of this restriction in m values for a photon is that its angular momentum vector is restricted to a unit value times Planck's constant parallel or anti-parallel to S .

The form of the modified spherical Hankel functions $\tilde{\xi}_l^\pm$ will depend on the functional form of $n(r)$. For example, let the special function for the index of refraction be given by

$$n^2 = 1 + \eta + \beta \left(\frac{r_o}{r} \right)^2 \quad (5.3-7)$$

where η and β are constants. This was introduced in [5]. From Eq. (5.3-5) it can be shown that this form offsets the spherical Hankel function in argument and spectral number:

$$\begin{cases} \tilde{\xi}_l^\pm(u) = \xi_l^\pm(\tilde{u}), \\ \tilde{u} = u\sqrt{1+\eta}, \quad u = kr \\ \tilde{l}^2 + \tilde{l} = (l^2 + l + \beta u_o^2) \end{cases} \quad (5.3-8)$$

For a thin atmosphere $\eta + \beta \approx 0$; these parameters may be individually chosen to match the index of refraction and its gradient at $u = u_o$. For example, for dry air at the Earth's surface $\eta + \beta \approx 1/4 \times 10^{-3}$ and $\beta \doteq -r_o n' \approx 0.2$. On the other hand, Eq. (5.3-7) does not satisfy our asymptotic boundary condition of $n(r) \rightarrow 1$ as $r \rightarrow \infty$. This form for $n(r)$ in Eq. (5.3-7) is useful for regional applications or over thin layers with boundaries on the top and bottom sides, and it has been used to study ducting, tunneling, super-refractivity, and other propagation effects in a strongly refracting medium.

Another technique, applicable when $n(r)$ assumes a general form, uses the WKB method to obtain an approximate expression for $\tilde{\xi}_l^\pm(u)$. We define $f_l(u)$ by

$$f_l = \frac{n^2 u^2 - l(l+1)}{u^2} \quad (5.3-9)$$

The WKB approximate solution $W_l^\pm(u) \doteq \tilde{\xi}_l^\pm(u)$, to Eq. (5.3-5) is given by [6]

$$W_l^\pm(u) = (f_l)^{-1/4} \exp \left(\pm i \int_{u_o}^u \sqrt{f_l} du \right) \quad (5.3-10)$$

Depending on the sign of $f_l(u)$, $W_l^\pm(u)$ has either an exponential form or a sinusoidal form. The WKB method has very widespread applicability. For examples in seismology see Chapman and Orcutt's review [1]. It also has been

mentioned in Chapter 4 in regard to wave propagation through a Cartesian stratified medium.

We will use an osculating parameter technique here. When $n(r)$ is variable, we may write

$$\tilde{\xi}_l^\pm(u) = a_l^\pm(\rho) \xi_l^\pm(\rho), \quad \rho = krn(r), \quad u = kr \quad (5.3-11)$$

where $a_l^\pm(\rho)$ is a so-called osculating parameter. It carries the deviation in amplitude and phase of $\tilde{\xi}_l^\pm(u)$ from these quantities in $\xi_l^\pm(\rho)$ due to the variability of $n(r)$.

The general series solution for the scalar potentials in a spherical stratified medium using this osculating parameter approach is given by

$$\begin{cases} \Pi^\pm(r, \theta, \phi) = \sum_{l=0}^{\infty} \left(a_l^\pm(\rho) \cos \phi + b_l^\pm(\rho) \sin \phi \right) \frac{\xi_l^\pm(\rho)}{\rho} P_l^1(\cos \theta), \\ \rho = krn(r) \end{cases} \quad (5.3-12)$$

In a homogeneous medium these spectral coefficients a_l^\pm and b_l^\pm are functions only of the spectral number, and their form depends on the asymptotic boundary conditions for the waveform; see Eqs. (3.2-4)-(3.2-6). In the inhomogeneous but spherical symmetric medium these spectral coefficients $a_l^\pm(\rho)$ and $b_l^\pm(\rho)$ vary also with ρ . The technique for obtaining their variability with ρ is rather similar to one of the parabolic equation techniques [7], but here their variability with ρ is due only to the gradient of the refractivity; the geometric component of the delay is retained by the $\xi_l^\pm(\rho)$ functions. Our task is to determine the form of these osculating spectral coefficients in a refracting medium in which a discontinuity also may be embedded, and to evaluate the series solutions for the electromagnetic field.

5.4 Asymptotic Forms

Because the spherical Bessel functions will be used extensively in later sections, we will need their asymptotic forms in terms of the Airy functions that are applicable for very large values of $\rho = knr$ and l . These have already been presented in Chapter 3, Section 3.8. There we established that the principal contributions to the spectral coefficients comes from spectral number values in the vicinity of $l = \rho$. Therefore, asymptotic forms that exploit the relatively small value of $|l - \rho|/\rho$ but the large value of ρ are appropriate. All of the asymptotic forms presented in Section 3.8 carry over to the stratified case here with the transformation $x = kr \rightarrow \rho = knr$ in the argument of the Bessel

functions, and with $y \rightarrow \hat{y}$ in the argument of the Airy functions. The argument \hat{y} is a function of $\rho = knr(r)$ and v . We have placed the caret over y to indicate its dependence on n through ρ . The key asymptotic forms used later are summarized here.

From Eqs. (3.2-8) and (3.8-9) we have for the spherical Hankel functions when $l \approx \rho^2$

$$\left. \begin{aligned} \xi_l^\pm(\rho) &\sim \sqrt{\pi K_\rho} \left(1 - K_\rho^{-2} \frac{\hat{y}}{15} + O[\rho^{-5/3}] \right) (\text{Ai}[\hat{y}] \mp i \text{Bi}[\hat{y}]) \\ \xi_l'^\pm(\rho) &\sim -\sqrt{\frac{\pi}{K_\rho}} \left(1 + K_\rho^{-2} \frac{\hat{y}}{15} + O[\rho^{-5/3}] \right) (\text{Ai}'[\hat{y}] \mp i \text{Bi}'[\hat{y}]) \end{aligned} \right\}, l \approx \rho \quad (5.4-1)$$

where $\rho = knr(r)$ and $v = l + 1/2$, and where $\text{Ai}[\hat{y}]$ and $\text{Bi}[\hat{y}]$ are the Airy functions of the first and second kind, respectively. See also [8]. For convenience we will use the spectral number l and $v = l + 1/2$ more or less interchangeably. The distinction between them is inconsequential because of the enormity of their values in the stationary phase neighborhoods. The argument \hat{y} is given by

$$\hat{y} = v^{2/3} \zeta(v/\rho) \quad (5.4-2)$$

Here the auxiliary function $\zeta(\mu)$, $\mu = v/\rho$, and its series expansions in powers of $[(\rho^2 - v^2)/v^2]$ and in powers of $[(\rho - v)/v]$ are defined in Eqs. (3.8-4) and (3.8-5) for both regimes $\mu \geq 1$ and $\mu \leq 1$. Using these expansions we summarize the key relationships between \hat{y} and ρ and v below:

$$\left. \begin{aligned} \hat{y} &= \frac{1}{4K_v^4} (v^2 - \rho^2) \left[1 + \frac{2}{5} \frac{v^2 - \rho^2}{v^2} + \dots \right] \\ \hat{y} &= \frac{1}{K_\rho} (v - \rho) \left(1 - \frac{v - \rho}{60K_\rho^3} + \dots \right) \\ v &= \rho + K_\rho \hat{y} \left(1 + \frac{\hat{y}}{60K_\rho^2} + \dots \right), \quad K_v = \left(\frac{v}{2} \right)^{1/3}, \quad K_\rho = \left(\frac{\rho}{2} \right)^{1/3} \end{aligned} \right\} \quad (5.4-3)$$

² When $\rho \gg l$ Eq. (5.4-1) is not appropriate. Starting from Eq. (3.8-10) it follows that $\xi_l^\pm(\rho) \rightarrow \mp i \exp(\pm iX)$, and $X = (\rho^2 - v^2)^{1/2} - v \cos^{-1}(v/\rho) + \pi/4 \rightarrow \rho - l\pi/2$. Hence $\xi_l^\pm(\rho) \rightarrow (\mp i)^{l+1} \exp(\pm i\rho)$ for $\rho \gg l$.

These truncated series expansions for \hat{y} and ν are very accurate for large values of ρ with $\nu \approx \rho$. For most stationary phase neighborhoods the value of \hat{y} will be small compared to K_ρ . Therefore, the term $K_\rho^{-2}\hat{y}/15$ in Eq. (5.4-1) can be dropped in the applications here. For GPS wavelengths $K_\rho^{-2}/15 \approx 3 \times 10^{-7}$.

The quantity $K_\rho = (\pi n r / \lambda)^{1/3}$, a quasi-constant, appears frequently throughout this monograph. For GPS wavelengths at sea level $K_\rho \approx 500$, and $2K_\rho / k \approx 30$ m. The latter turns out to be the spatial distance over which the Airy functions asymptotically transform from exponential functions to sinusoidal functions.

We also will need the asymptotic forms for the Airy functions. See [8] for a comprehensive discussion. They also are given in Chapter 3, Eq. (3.8-7) applicable for negative values of \hat{y} and also by Eq. (3.8-8) for positive values.

5.5 Modified Mie Scattering in a Spherical Stratified Medium

The central task in this section is to derive the spectral density function for the phase delay incurred by the l th spectral component of the wave as a result of the refractive gradient of the medium. This function $G[\rho, \nu]$, with $\nu = l + 1/2$, accounts for the extra phase delay in the l th spectral coefficient induced by only the refractive gradient of the medium. The geometric component of the phase delay is carried by the spherical Hankel function.

To follow a Mie scattering approach, we use the scalar potentials for the approaching, transmitted and scattered wave. Electric and magnetic scalar potentials, ${}^e\Pi^{(i)}$ and ${}^m\Pi^{(i)}$, were discussed in Chapter 3 and also in Section 5.2. An incoming planar harmonic wave with in-plane polarization and with zero phase at $\theta = \pi/2$ can be represented by series solutions in terms of spherical Bessel functions and spherical harmonic functions. For a non-conducting homogeneous medium these representations are given by

$$\left. \begin{aligned} {}^e\Pi_l &= \frac{E_0}{nk} i^{l-1} \frac{(2l+1)}{l(l+1)} \frac{\psi_l(nu)}{nu} P_l^1(\cos\theta) \cos\phi \\ {}^m\Pi_l &= \frac{H_0}{nk} i^{l-1} \frac{(2l+1)}{l(l+1)} \frac{\psi_l(nu)}{nu} P_l^1(\cos\theta) \sin\phi \\ E_0 \sqrt{\epsilon} &= H_0 \sqrt{\mu}; \quad u = kr, \quad n = \text{constant} \end{aligned} \right\} \quad (5.5-1)$$

Here ${}^e\Pi_l$ and ${}^m\Pi_l$ are the l th spectral components of the electric and magnetic scalar potentials, respectively.

To obtain the electromagnetic field from these scalar potentials, one uses the vector curl operations on their vector form given in Eq. (5.2-7). This vector form, $(\Pi_l \mathbf{r})$, is known as the Hertz potential. Here E_o is the amplitude of the electric field vector that lies in the plane defined by $\phi = 0$, that is, along the \hat{x} direction in Figure 4-10. Similarly, H_o is the amplitude of the magnetic field vector, which points in the \hat{y} direction. From Maxwell's equations it follows that $E_o \sqrt{\epsilon} = H_o \sqrt{\mu}$.

Following that treatment for the homogeneous case, we obtain the series expansion solutions for the scalar potentials of the incoming wave in the spherical symmetric stratified medium with an index of refraction $n(u)$. Here the scalar potentials are given by

$$\left. \begin{aligned} {}^e \tilde{\Pi} &= \frac{E_o}{nk} \sum_{l=1}^{\infty} {}^e a_l(nu) \frac{\psi_l(nu)}{nu} P_l^1(\cos \theta) \cos \phi \\ {}^m \tilde{\Pi} &= \frac{H_o}{nk} \sum_{l=1}^{\infty} {}^m a_l(nu) \frac{\psi_l(nu)}{nu} P_l^1(\cos \theta) \sin \phi \\ E_o \sqrt{\epsilon} &= H_o \sqrt{\mu}, \quad u = kr \end{aligned} \right\} \quad (5.5-2)$$

The main difference from the homogeneous case is that we have introduced the spectral coefficients $a_l(nu)$, which are now variable with u , to account for the effects of the variability in $n(r)$. Each spectral component of these series satisfies the modified Helmholtz equation in Eq. (5.2-8); thus, the product $a_l(nu)\psi_l(nu)$ constitutes a formal solution to the modified spherical Bessel equation in Eq. (5.3-3). For each integer value of l , $a_l(nu)$ is an osculating parameter. The osculating parameter technique has been already discussed in Section 4.8 for a Cartesian-stratified medium. This technique is useful for solving certain ordinary differential equations where the rapidly varying component is carried by the basis function, $\psi_l(nu)$ in this case, and the more slowly (sometimes) varying component is carried by $a_l(nu)$.

We will need the asymptotic form for $a_l(nu)$ corresponding to an incoming wave well outside the atmosphere and its refractivity or scattering effects. The asymptotic form depends on where we place the emitting GPS satellite, either a finite or an infinite distance away, but always in the direction $\theta = \pi$. For the infinite case the incoming waves are planar, and it follows from Eq. (5.5-1) that $a_l(nu)$ has the limit

$$a_l(nu) \xrightarrow[u \rightarrow \infty, n \rightarrow 1]{} i^{l-1} \frac{2l+1}{l(l+1)} \quad (5.5-3a)$$

This form satisfies the asymptotic boundary condition that the approaching wave must be planar at large distances and travelling in the direction of the positive z -axis (see Eq. (4.11-1) and Figure 4-10). The form of the approaching wave is $\exp(iu_L \cos \theta_L)$. This is referenced to the phase on the line $\theta = \pi/2$.

For the case of the GPS satellite at a finite distance we have to account for the arrival of a spherical wave, with its center at the transmitting GPS satellite instead of a planar or collimated wave. Referring to Figure A-3 in Appendix A, this spherical wave is given by $\exp(iu_{LG})/u_{LG}$. In this case the asymptotic form for $a_l(nu)$ is more complicated than that given in Eq. (5.5-3a) because it must correspond to the spectral component of the spherical waveform, which explicitly includes the location of the transmitter. From Section 5.3 where Bauer's identity is derived from the multipole expansion for a spherical wave, one can work out the correct asymptotic form for the spherical case. It is given by³

$$a_l(nu) \rightarrow A i^{l-1} \frac{2l+1}{l(l+1)} \left(i^{l+1} \frac{\xi_l^+(u_G)}{u_G} \right) \quad (5.5-3b)$$

Here the phase in this asymptotic form is now referenced to the position r_G of the transmitting GPS satellite⁴. The amplitude A is a constant. For example, if

³ Eq. (5.5-3b) follows from the homogeneous case, $n \equiv 1$, by first noting that

$$\frac{i}{u_L} \frac{\partial}{\partial \theta_L} \left(\frac{\exp[iu_{LG}]}{u_{LG}} \right) = \left(\frac{\exp[iu_{LG}]}{u_{LG}} \right) \left(1 + \frac{i}{u_{LG}} \right) \sin(\theta_L - \chi_G)$$

where χ_G is the deflection angle of the straight line between the transmitting GPS satellite and the LEO (see Figure A-3). The GPS satellite is located at (r_G, θ_G) , but always in the direction $\theta_G = \pi$. The radial component of the electric field at the LEO from the spherical wave centered at (r_G, θ_G) is given by

$$E_r(r_L, \theta_L) = \left(\frac{\exp[iu_{LG}]}{u_{LG}} \right) \sin(\theta_L - \chi_G)$$

Dropping the i/u_{LG} term, it follows that

$$E_r(r_L, \theta_L) \doteq \frac{i}{u_L} \frac{\partial}{\partial \theta} \left(\frac{\exp[iu_{LG}]}{u_{LG}} \right) = \sum_{l=0}^{\infty} (-1)^l (2l+1) \frac{\psi_l(u_L)}{u_L^2} \frac{\xi_l^+(u_G)}{u_G} P_l(\cos \theta_L)$$

The RHS of this equation comes from the multipole expansion for $(\exp[iu_{LG}]/u_{LG})$ given in Eq.(5.3-2). Equating this series form for $E_r(r_L, \theta_L)$ to the form obtained from the corresponding vector calculus operations on the trial scalar potential series (see Eqs. (5.5-7) and (5.5-8)) yields the asymptotic form for the spectral coefficients given Eq. (5.5-3b). Getting the coefficients for one component of the field, $E_r(r_L, \theta_L)$ in this case, is sufficient.

⁴ We can use the asymptotic form

we renormalize the amplitude by setting $A = u_{LG}$, then in the limit as $r_G, r_{LG} \rightarrow \infty$, the asymptotic form for $a_l(nu)$ in Eq. (5.53b), but referenced to the $\theta = \pi/2$ line, approaches the form given in Eq. (5.5-3a) for the collimated wave. In any case, we will assume the collimated form in Eq. (5.5-3a) subsequently. The correction for the case of an incident spherical wave appears straightforward, and it is noted in Section 5.10 and Chapter 6.

To develop a functional form for $a_l(nu)$ we first will obtain the change in $a_l(nu)$ that results from a change in the index of refraction across a spherical boundary, which is embedded in an *otherwise homogeneous* medium and located at $r = r_o$. By applying the continuity conditions from Maxwell's equations, the spectral coefficients for the transmitted and reflected waves are expressed in terms of the spectral coefficients of the incident wave at the boundary and the change in refractivity. After obtaining the changes in the spectral coefficients that apply across a boundary we will use a limiting procedure to obtain a continuous version for these spectral coefficients.

The change in $a_l(nu)$ obtained in this manner is characterized by a first order differential equation. On the other hand, Maxwell's equations comprise a second order system for this essentially two-dimensional problem. (See Section 4.11) Therefore, this approach involves an approximation, the accuracy of which we will establish. We saw in the Cartesian case discussed in Sections 4.8 and 4.9 that this approximation works well for points sufficiently distant from a turning point and when thin atmosphere conditions apply. The same conclusions hold here, although the concept of a turning point in a wave theory approach has to be expressed in terms of both the radial coordinate ρ and the spectral number l .

Chapter 3, Sections 3.3 and 3.5, and also Chapter 4, Section 4.6 for the Cartesian case, discuss the formalism for treating standing electromagnetic waves in terms of a spectral composition of incoming and outgoing waves. In Chapter 3 the spherical Bessel function was bifurcated into the spherical Hankel function of the first kind to represent outgoing waves and its equally weighted complex conjugate, the spherical Hankel function of the second kind, to represent incoming waves. Specifically, the spherical Hankel functions of the

$$i^{l+1} \xi_l^+(u_G) \rightarrow (u_G^2 / (u_G^2 - v^2))^{1/4} \exp[i(\sqrt{u_G^2 - v^2} + v \sin^{-1}(v/u_G))], \quad v = l + 1/2$$

in Eq. (5.5-3b) because u_G will be very much larger than the range of spectral numbers yielding stationary values for the spectral series. If the phase terms here are added to the spectral density function for the phase delay through the atmosphere given from the collimated case, we have the correct form for the phase for the case where the incident wave is spherical. See Chapter 5.10, Eq. (5.10-12). The term $(u_G^2 / (u_G^2 - v^2))^{1/4}$ is related to the reduced limb distance used to convert the geometry with u_G finite to an equivalent geometry with u_G infinite.

first (+) and second (−) kinds, ξ_l^\pm , are defined by $\xi_l^\pm = \psi_l \pm i\chi_l$, where $\psi_l(x) = (\pi x/2)^{1/2} J_{l+1/2}(x)$ and $\chi_l(x) = (\pi x/2)^{1/2} Y_{l+1/2}(x)$, where $J_l(x)$ and $Y_l(x)$ are the integer Bessel functions of the first and second kind, respectively. Using the asymptotic forms for the Bessel functions applicable when $x \gg l$, one can readily show that ξ_l^+ assumes the form that describes an outgoing spherical wave, and that ξ_l^- describes an incoming spherical wave. In a homogeneous medium outgoing waves interior to the scattering boundary are generated from incoming waves that reflect around the origin, which the scattering coefficients b_l show as $r \rightarrow 0$. This formalism was necessary to treat internal reflections at the boundary of the scattering sphere and to isolate the scattering coefficients for an emerging wave that has undergone a specific number of internal reflections.

We adopt the same formalism here. Thus, the electric field at any point will be treated as a spectral composition of radial incoming and radial outgoing wavelets, which are combined in a weighted summation over all spectral numbers. They also are combined in such a way to eliminate the singularity at the origin arising from the Bessel function of the second kind.

5.5.1 Incoming Waves

Let us first consider an incoming incident wave. Here the scalar potentials (see Eq. (5.5-1)) that generate $\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$ are given by

$$\left. \begin{aligned} {}^e\Pi^{(i)} &= \frac{E_o}{2n_1k} \sum_{l=1}^{\infty} {}^e a_l^{(i)} \frac{\xi_l^-(n_1u)}{n_1u} P_l^1(\cos\theta) \cos\phi \\ {}^m\Pi^{(i)} &= \frac{H_o}{2n_1k} \sum_{l=1}^{\infty} {}^m a_l^{(i)} \frac{\xi_l^-(n_1u)}{n_1u} P_l^1(\cos\theta) \sin\phi \end{aligned} \right\} \quad (5.5-4)$$

Here $u = kr$ is the radial coordinate expressed in phase units. The scattering boundary is located at u_o ; ε_1 and μ_1 are constants that define the index of refraction in the homogeneous medium on the incident side of the boundary where $u \geq u_o$; $a_l^{(i)}$ is spectral coefficient for the incoming incident wave. Because $\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$ are the fields for an incoming wave at the boundary we must use the spherical Hankel functions of the second kind $\xi_l^-/2$ for the radial function instead of ψ_l for determining the spectral coefficients at the boundary⁵.

⁵ Recall that $\psi_l = (\xi_l^+ + \xi_l^-)/2$. If we did use ψ_l in the scalar potential series for the incoming incident wave, we would find upon applying the continuity conditions at the

Similarly, the scalar potentials for the scattered or reflected wave are given in terms of the scattering coefficients b_l by

$$\left. \begin{aligned} {}^e\Pi^{(S)} &= \frac{E_o}{n_1 k} \sum_{l=1}^{\infty} {}^e b_l \frac{\xi_l^+(n_1 u)}{n_1 u} P_l^1(\cos\theta) \cos\phi \\ {}^m\Pi^{(S)} &= \frac{H_o}{n_1 k} \sum_{l=1}^{\infty} {}^m b_l \frac{\xi_l^+(n_1 u)}{n_1 u} P_l^1(\cos\theta) \sin\phi \end{aligned} \right\} \quad (5.5-5)$$

Because the scattered wave is outgoing we must use the spherical Hankel functions of the first kind ξ_l^+ in this representation in order to match the asymptotic boundary condition as $r \rightarrow \infty$, which requires a spherical wavefront from a scattering surface (and which the ξ_l^+ function indeed provides in its asymptotic form for large r). Finally, the scalar potentials for the transmitted wave, which is incoming, are given by

$$\left. \begin{aligned} {}^e\Pi^{(T)} &= \frac{E_o}{2n_2 k} \sum_{l=1}^{\infty} {}^e a_l^T \frac{\xi_l^-(n_2 u)}{n_2 u} P_l^1(\cos\theta) \cos\phi \\ {}^m\Pi^{(T)} &= \frac{H_o}{2n_2 k} \sum_{l=1}^{\infty} {}^m a_l^T \frac{\xi_l^-(n_2 u)}{n_2 u} P_l^1(\cos\theta) \sin\phi \end{aligned} \right\} \quad (5.5-6)$$

Here $a_l^{(T)}$ is the spectral coefficient for the wave transmitted across the boundary located at u_o ; ε_2 and μ_2 are constants that define the index of refraction on the transmitted side of the boundary where $u \leq u_o$.

To obtain the continuity conditions consider first the electromagnetic field generated by the scalar potential ${}^e\Pi(r, \theta, \phi)$, which generates the TM wave. From Eq. (5.2-7) and using the identity $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, one obtains

$$\mathbf{E}_{TM} = (2 + \mathbf{r} \cdot \nabla) \nabla({}^e\Pi) - \mathbf{r} \nabla^2({}^e\Pi), \quad \mathbf{H}_{TM} = -ik\varepsilon \nabla({}^e\Pi) \times \mathbf{r} \quad (5.5-7)$$

boundary that the scattering coefficients $b_l(nu)$ would carry an extra "-1" term that would exactly cancel the $\xi_l^+ / 2$ part of ψ_l , effectively leaving only the $\xi_l^- / 2$ part to represent the entire field, incident plus scattered. For this case where ψ_l is used in the incident series, as $(n_2 - n_1)u \rightarrow 0$, $b_l(nu) \rightarrow -a_l^{(i)}(nu) / 2$.

For example, using Eq. (5.5-4) for the incident wave, (and using the differential equation for the spherical Bessel function, $d^2 \xi_l / d\rho^2 + (1 - l(l+1) / \rho^2) \xi_l = 0$, for E_r) the field components in Eq. (5.5-7) become

$$\left. \begin{aligned} [E_r^{(i)}]_{\text{TM}} &= \left(\frac{\partial^2}{\partial r^2} ({}^e \Pi^{(i)} r) + n_1^2 k^2 {}^e \Pi^{(i)} r \right) \\ &= \frac{E_o^{(i)}}{n_1 u} \sum_{l=1}^{\infty} l(l+1) {}^e a_l^{(i)} \frac{\xi_l^-(n_1 u)}{2n_1 u} P_l^1 \cos \phi, \\ [E_\theta^{(i)}]_{\text{TM}} &= \frac{1}{r} \frac{\partial^2}{\partial r \partial \theta} ({}^e \Pi^{(i)} r) = E_o^{(i)} \sum_{l=1}^{\infty} {}^e a_l^{(i)} \frac{\xi_l^-(n_1 u)}{2n_1 u} \frac{dP_l^1}{d\theta} \cos \phi, \\ [E_\phi^{(i)}]_{\text{TM}} &= \frac{1}{r} \frac{\partial^2}{\partial r \partial \phi} ({}^e \Pi^{(i)} r) = -E_o^{(i)} \sum_{l=1}^{\infty} {}^e a_l^{(i)} \frac{\xi_l^-(n_1 u)}{2n_1 u} P_l^1 \sin \phi, \\ [H_r^{(i)}]_{\text{TM}} &\equiv 0, \\ [H_\theta^{(i)}]_{\text{TM}} &= \frac{k\varepsilon}{ir} \frac{\partial}{\partial \phi} ({}^e \Pi^{(i)} r) = \frac{iH_o^{(i)}}{\sin \theta} \sum_{l=1}^{\infty} {}^e a_l^{(i)} \frac{\xi_l^-(n_1 u)}{2n_1 u} P_l^1 \sin \phi, \\ [H_\phi^{(i)}]_{\text{TM}} &= \frac{ik\varepsilon}{r} \frac{\partial}{\partial \theta} ({}^e \Pi^{(i)} r) = iH_o^{(i)} \sum_{l=1}^{\infty} {}^e a_l^{(i)} \frac{\xi_l^-(n_1 u)}{2n_1 u} \frac{dP_l^1}{d\theta} \cos \phi, \\ E_o^{(i)} \sqrt{\varepsilon_1} &= H_o^{(i)} \sqrt{\mu_1}, \quad P_l^1 = P_l^1(\cos \theta) \end{aligned} \right\} \quad (5.5-8)$$

We can write a set of expressions of a similar form for the scattered fields, $\mathbf{E}^{(S)}$ and $\mathbf{H}^{(S)}$, and for the transmitted fields, $\mathbf{E}^{(T)}$ and $\mathbf{H}^{(T)}$. Using the symmetry properties of the electromagnetic field discussed in Section 3.2, one also can readily develop a set of expressions from ${}^m \Pi$ for the TE wave. The complete field is given by the sum of these TM and TE expressions.

To obtain the required relationships between the spectral coefficients we use the continuity conditions from Maxwell's equations that the various field components must satisfy. Across a boundary with neither surface charges nor surface currents, Maxwell's equations require the components of the electromagnetic field to satisfy the following continuity conditions

$$\left. \begin{aligned} \mathbf{E}_{\tan g}^{(i)} + \mathbf{E}_{\tan g}^{(S)} &= \mathbf{E}_{\tan g}^{(T)}, \quad \varepsilon_1 (\mathbf{E}_r^{(i)} + \mathbf{E}_r^{(S)}) = \varepsilon_2 \mathbf{E}_r^{(T)} \\ \mathbf{H}_{\tan g}^{(i)} + \mathbf{H}_{\tan g}^{(S)} &= \mathbf{H}_{\tan g}^{(T)}, \quad \mu_1 (\mathbf{H}_r^{(i)} + \mathbf{H}_r^{(S)}) = \mu_2 \mathbf{H}_r^{(T)} \end{aligned} \right\} \quad (5.5-9a)$$

Here ε_1 and μ_1 apply to the incident side of the boundary; ε_2 and μ_2 apply to the transmitted side.

We apply these continuity conditions to the vector fields generated from the scalar potentials in Eqs. (5.5-4), (5.5-5) and (5.5-6) for a boundary located at $r = r_o$. The electromagnetic field for the incident TM wave is shown in Eq. (5.5-8), but because they are all similar, we forego writing the other five sets for the scattered and transmitted TM waves and for all TE modes. Applying the continuity conditions in Eq. (5.5-9a) to these waves at all applicable points on the boundary of the sphere located at $r = r_o$, we obtain an equivalent set of continuity conditions that only involve the individual spectral coefficients and their Hankel functions. These conditions written in matrix form become

$$\left\{ \begin{array}{l} \begin{pmatrix} \frac{\xi_l^-(n_2 u_o)}{\mu_2} & -\frac{\xi_l^+(n_1 u_o)}{\mu_1} \\ \frac{\xi_l'^-(n_2 u_o)}{n_2} & -\frac{\xi_l'^+(n_1 u_o)}{n_1} \end{pmatrix} \begin{pmatrix} {}^e a_l^{(T)} \\ 2 {}^e b_l \end{pmatrix} = \begin{pmatrix} \frac{\xi_l^-(n_1 u_o)}{\mu_1} \\ \frac{\xi_l'^-(n_1 u_o)}{n_1} \end{pmatrix} {}^e a_l^{(i)} \\ \begin{pmatrix} \frac{\xi_l^-(n_2 u_o)}{n_2} & -\frac{\xi_l^+(n_1 u_o)}{n_1} \\ \frac{\xi_l'^-(n_2 u_o)}{\mu_2} & -\frac{\xi_l'^+(n_1 u_o)}{\mu_1} \end{pmatrix} \begin{pmatrix} {}^m a_l^{(T)} \\ 2 {}^m b_l \end{pmatrix} = \begin{pmatrix} \frac{\xi_l^-(n_1 u_o)}{n_1} \\ \frac{\xi_l'^-(n_1 u_o)}{\mu_1} \end{pmatrix} {}^m a_l^{(i)} \end{array} \right\}, l = 1, 2, \dots \quad (5.5-9b)$$

Solving the linear system of equations in Eq. (5.5-9b) for the transmission and scattering coefficients in terms of the incident coefficients, we obtain

$$\left\{ \begin{array}{l} {}^e a_l^{(T)} = \left[\frac{-2i}{n_1 \mu_1 {}^e \mathcal{W}_l'} \right] {}^e a_l^{(i)}, \quad 2 {}^e b_l = - \left[\frac{{}^e \mathcal{W}_l^-}{{}^e \mathcal{W}_l'} \right] {}^e a_l^{(i)} \\ {}^m a_l^{(T)} = \left[\frac{-2i}{n_1 \mu_1 {}^m \mathcal{W}_l'} \right] {}^m a_l^{(i)}, \quad 2 {}^m b_l = - \left[\frac{{}^m \mathcal{W}_l^-}{{}^m \mathcal{W}_l'} \right] {}^m a_l^{(i)} \end{array} \right\}, l = 1, 2, \dots \quad (5.5-10)$$

where

$$\left\{ \begin{array}{l} {}^e\mathcal{W}_l = \frac{\xi_l^+(n_1 u_o) \xi_l'^-(n_2 u_o)}{n_2 \mu_1} - \frac{\xi_l'^+(n_1 u_o) \xi_l^-(n_2 u_o)}{n_1 \mu_2} \\ {}^e\mathcal{W}_l^\pm = \frac{\xi_l^\pm(n_1 u_o) \xi_l'^\pm(n_2 u_o)}{n_2 \mu_1} - \frac{\xi_l'^\pm(n_1 u_o) \xi_l^\pm(n_2 u_o)}{n_1 \mu_2} \\ {}^m\mathcal{W}_l = \frac{\xi_l^+(n_1 u_o) \xi_l'^-(n_2 u_o)}{n_1 \mu_2} - \frac{\xi_l'^+(n_1 u_o) \xi_l^-(n_2 u_o)}{n_2 \mu_1} \\ {}^m\mathcal{W}_l^\pm = \frac{\xi_l^\pm(n_1 u_o) \xi_l'^\pm(n_2 u_o)}{n_1 \mu_2} - \frac{\xi_l'^\pm(n_1 u_o) \xi_l^\pm(n_2 u_o)}{n_2 \mu_1} \end{array} \right\}, l = 1, 2, \dots \quad (5.5-11)$$

The Wronskian of the spherical Hankel functions

$$\mathcal{W}[\xi_l^+(z), \xi_l^-(z)] = \xi_l^+ \xi_l'^- - \xi_l'^+ \xi_l^- = -2i \quad (5.5-12)$$

has been used in Eq. (5.5-9b) to obtain the transmission coefficients in Eq. (5.5-10).

The “electric” coefficients $({}^e a_l, {}^e b_l)$ and the “magnetic” coefficients $({}^m a_l, {}^m b_l)$ differ from their counterparts by a small quantity of the order $N = n - 1$. Because we have assumed a thin atmosphere $N(r) \ll 1$; we will ignore this difference herewith and in the interest of simplifying the notation, we will suppress the superscripts “e” and “m” on the scattering coefficients and retain only the electric coefficients in the following. These small differences can readily be reconstituted to obtain the scattered wave from the vector calculus operations on both the electric and magnetic scalar potentials. Also, for the case where \mathbf{E} lies in the plane $\phi = 0$, one can show that for large spectral numbers the magnetic coefficients provide a negligible contribution to the field. This follows from noting that the magnetic coefficients involve $P_l^1(\cos\theta)$, whereas the electric coefficients involve $dP_l^1/d\theta$. However, $P_l^1/(dP_l^1/d\theta) \sim l^{-1} \ll 1$.

We note that the GPS signals are principally right-hand circular polarized; therefore, to study polarization effects from the refracting sphere, we would need to retain the cross-plane polarization ($\phi = \pi/2$) scattering terms also, which are appropriately offset in phase to secure the proper elliptical or circular polarization. However, for $N \ll 1$ the scattering for the two linear polarization modes differ by an amount of the order N . Also, because of the previously mentioned relativistic covariance of the electrodynamics equations, we can exploit that symmetry to convert the solution for $\mathbf{H}^{(S)}$ for the in-plane

polarized case discussed here directly into a solution for $E^{(S)}$ for the cross-plane polarization case.

For outgoing waves, for example, for waves that have passed through the scattering sphere or, in a geometric optics context, rays that have passed their point of tangency with an arbitrary spherical boundary at radius $r = r_*$, one would obtain an analogous system of transfer equations to those given in Eq. (5.5-10). The only difference is that the scalar potential series for the incident and transmitted waves would each carry the $\xi_l^+ / 2$ functions instead of the $\xi_l^- / 2$ functions because they are outgoing. Also, the scalar potential series for the waves reflected from the inner side of the boundary would carry the $\xi_l^- / 2$ functions because they are incoming after being reflected.

5.5.2 Evaluating the Spectral Coefficients in a Stratified Medium

We now set $\mu(r) \equiv 1$ in the following discussion, which further simplifies the notation, albeit at the price of losing the symmetries in Eqs. (5.5-10) and (5.5-11).

Next, we treat the continuously varying refractivity in the medium as a series of concentric shells. Within each spherical shell the refractivity is a constant, but it changes discontinuously across the boundary of each shell. So, the refractivity varies in the radial direction in a stepwise manner. This is the thin film model, or one version of the so-called onion skin model. Across each boundary the transition equations for the spectral coefficients in Eq. (5.5-10) apply. After obtaining these spectral coefficients across the boundary of each shell, we will let the number of shells grow infinite while requiring their individual widths to become infinitesimal in such a way that the ensemble spans the appropriate physical space or range.

At the boundary located at $u_o = kr_o$ we let $n_1 = n - \Delta n / 2$ and $n_2 = n + \Delta n / 2$ where Δn is sufficiently small that $u\Delta n$ can be considered as an infinitesimal. Expanding $n_1 n_2 \mathcal{W}_l$ and $n_1 n_2 \mathcal{W}_l^\pm$ in powers of $u\Delta n$, we obtain

$$\left. \begin{aligned} n\mathcal{W}_l &= -2i - \frac{\Delta n}{2n} \left[nu \left(2\xi_l^{+'} \xi_l^{-'} - \xi_l^{-} \xi_l^{+''} - \xi_l^{+} \xi_l^{-''} \right) \right. \\ &\quad \left. + \xi_l^{+'} \xi_l^{-} + \xi_l^{+} \xi_l^{-'} \right]_{nu_o} + O[(u\Delta n)^2], \\ n\mathcal{W}_l^\pm &= -\frac{\Delta n}{n} \left[nu \left(\xi_l^{\pm'} \xi_l^{\pm'} - \xi_l^{\pm} \xi_l^{\pm''} \right) + \xi_l^{\pm'} \xi_l^{\pm} \right]_{nu_o} + O[(u\Delta n)^2] \end{aligned} \right\} \quad (5.5-13)$$

It follows that as $\Delta n \rightarrow 0$, $\mathcal{W}_l^\pm \rightarrow 0$, but $n\mathcal{W}_l \rightarrow -2i$. From Eq. (5.5-11) it follows that $b_l \rightarrow 0$ and that $a_l^{(T)} \rightarrow a_l^{(i)}$ when $\Delta n \rightarrow 0$.

For a series of concatenated shells, multiple internal reflections should be considered. For example, outward reflected rays from inner shell boundaries will again be reflected inward at the boundary of interest. We have already discussed this in Section 4.8 for Cartesian layers, and Figure 4-8 in that section applies as well here. Specifically, we can use the discussion in Section 4.12 to transform our spherical geometry here into an equivalent Cartesian-stratified geometry involving Airy layers. By this means, conclusions drawn from the Cartesian case can be applied here. In Section 4.8 we showed that the ensemble of doubly reflected rays that add to the incident wave each involve a factor of the order of Δn^2 (here Δn is the average change in index of refraction from layer to layer). Moreover, the phase of these secondary rays (at the right-hand boundary of the j th layer in Figure 4-8) will be randomly distributed when the span Δr of the layers is such that $\Delta r \gg \lambda$. It can be shown by vector summing up the contributions from all of these reflected rays with a second reflection from the left-hand boundary of the j th layer, that the ratio of their combined contributions to the main ray contribution is given by $n'\lambda$, which is negligible for a thin atmosphere. Therefore, in calculating the spectral coefficients for the transmitted wave through a transparent medium we can neglect secondary and higher order reflections in our shell model when thin atmosphere conditions apply and provided that we avoid turning points.

The incident field at the $j+1$ st boundary can be considered as the product of the transmission coefficients from the previous j layers. If we then expand that product and retain only the first order terms, we can obtain a first order differential equation for the spectral coefficients. The range of validity of this linear truncation is essentially the same as that found for the truncation of the characteristic matrix to linear terms given in Section 4.4. There we found for a thin atmosphere that the accuracy of this truncation was satisfactory provided that we stay clear of turning points.

Let us define $(a_l^-)_j$ to be the l th spectral coefficient of an incoming transmitted wave for the j th layer. The superscript “-“ on a_l^- denotes an incoming wave. We drop herewith the superscripts “ i ” and “ T ”. Then, using Eqs. (5.5-10) and (5.5-13), it follows that

$$(a_l^-)_{j+1} \doteq (a_l^-)_j \left[\frac{1 + \Delta n / 2n}{1 - i g_l \Delta n / n} \right]_j \quad (5.5-14)$$

where $g_l(\rho_j)$ is a function of the spherical Hankel functions obtained from Eq. (5.5-13), which is defined in Eq. (5.5-19) and will be discussed shortly. Here we define $\rho = un(r) = krn(r)$. For a series of layers it follows from Eq. (5.5-14) that

$$(a_l^-)_{k+1} \doteq (a_l^-)_1 \prod_{j=1}^k \left(\frac{1 + \frac{\Delta n_j}{2n_j}}{1 - ig_l(\rho_j) \frac{\Delta n_j}{n_j}} \right) \quad (5.5-15)$$

To evaluate Eq. (5.5-15) we note that $\log[\prod(*)] = \sum[\log(*)]_j$. When $g_l \Delta n_j \ll 1$, we can expand $\log[1 - ig_l(\rho_j)(\Delta n_j / n_j)]$, retaining only first order terms in Δn_j . Thus, Eq. (5.5-15) becomes

$$\log \left(\frac{(a_l^-)_{k+1}}{(a_l^-)_1} \right) \doteq \sum_{j=1}^k \left(\left(\frac{1}{2} + ig_l(\rho_j) \right) \frac{\Delta n_j}{n_j} \right) \quad (5.5-16)$$

We set⁶ $\Delta n = (dn / d\rho) \Delta \rho$. Also, we define $\Delta n_j = n_{j+1} - n_j$ to be the change in the index of refraction across the j th boundary (Figure 4-8), and we define $\Delta \rho_j = \rho_{j+1} - \rho_j$ to be the optical thickness of the j th layer. From Eq. (5.5-16) it follows that in the limit as $\Delta \rho \rightarrow 0$, we obtain

$$\frac{1}{a_l^-} \frac{da_l^-}{d\rho} = \left(\frac{1}{2} + ig_l(\rho) \right) \frac{d \log n}{d\rho} \quad (5.5-17)$$

Here $g_l(\rho)$ is defined by

$$g_l(\rho) = \left[\frac{\rho}{2} (\xi_l^{\pm'} \xi_l^{\mp'} - \xi_l^{\pm} \xi_l^{\mp}) + \frac{1}{4} (\xi_l^+ \xi_l^{-'} + \xi_l^{+'} \xi_l^-) \right]_{\rho=un} \quad (5.5-18)$$

Bessel's equation in Eq. (5.3-3) has implicitly been used in Eq. (5.5-18). The enormity of $\rho \sim 10^8$ allows us to ignore the second term $(\xi_l^+ \xi_l^{-'} + \xi_l^{+'} \xi_l^-) / 4$. Using Bessel's equation to replace $\xi_l^{\pm''}$ and dropping the relatively small term in Eq. (5.5-18), one obtains

⁶Note that $d \log n / d\rho = (n + u dn / du)^{-1} d \log n / du$. Also, $\rho(d \log n / d\rho) = u(d \log n / du) / (1 + u(d \log n / du))$. The quantity $u \mid d \log n / du \mid$ is the ratio of the radius of curvature (r) of the spherical boundary to the local radius of curvature of the ray ($n \mid dn / du \mid$). It is the parameter β defined in Chapter 2, Eq. (2.2-9), which is small for a thin atmosphere (for dry air in the Earth's atmosphere at sea level this ratio is about 0.2). In a super-refracting medium, occasionally caused by a water vapor layer in the lower troposphere, $d\rho / du < 0$. Across a boundary $d\rho / du = 0$, which requires reverting to the variable u in Eq. (5.5-17).

$$g_l(\rho) \doteq \frac{\rho}{2} \left(\xi_l^+ \xi_l'^- + \left(1 - \frac{l(l+1)}{\rho^2} \right) \xi_l^+ \xi_l^- \right) \quad (5.5-19)$$

Figure 5-2 shows $g_l(\rho)$ versus \hat{y} including its asymptotic forms. Here \hat{y} is the argument of the Airy functions. The relationship between \hat{y} , l and $\rho = un$ was discussed earlier in Section 5.4, Eqs. (5.4-2) and (5.4-3). It suffices here to note that $v = l + 1/2 \doteq \rho + \hat{y}(\rho/2)^{1/3}$ to very high accuracy when ρ is large and \hat{y} is relatively small. For \hat{y} values greater than about +2, $g_l(\rho)$ is dominated by the spherical Bessel function of the second kind and it breaks sharply to very large negative values.

The derivation for $da_l^-/d\rho$ fails for this regime, $\hat{y} > 0$, because the basic assumption that $g_l \Delta n_j \ll 1$ in Eq. (5.5-16) is invalid when $g_l(\rho) \rightarrow \infty$ for increasing $v > \rho$. In fact, the correct form for $g_l(\rho)$ rapidly approaches zero for $v > \rho$, rather than blowing up, as the form for $g_l(\rho)$ given in Eq. (5.5-19) does. The modified Mie scattering derivation that we have used did not account for curvature terms and it assumes that $g_l(\rho) \Delta n$ can be made a small quantity, which is not valid below a turning point. We return this issue in Section 5.7, after a discussion of asymptotic forms. There we present one method for asymptotic matching the $g_l(\rho)$ function given in Eq. (5.5-19) with a version that does hold for $\hat{y} > 0$.

In general the initial condition for a_l^- in Eq. (5.5-17) depends on the boundary conditions for the electromagnetic field. In a geometric optics context, the initial condition for a_l^- is a ray-specific quantity, that is, it depends at least in part on the impact parameter of the ray (or cophasal normal path) associated with the wave as it propagates through the medium. Therefore, the constant of integration obtained from integrating Eq. (5.5-17) will depend on ray-specific boundary conditions. However, in the special case where the approaching rays are collimated before encountering the medium, they all have the same asymptotic boundary condition as $u \rightarrow \infty$; in this case the constant of integration will be invariant with impact parameter. For departing waves this symmetry is spoiled⁷ by the intervening refracting medium, and the asymptotic boundary conditions as $u \rightarrow \infty$ will vary with the impact parameter of the approaching ray.

⁷ We could, however, form a symmetric problem merely by forcing the electromagnetic wave to be planar along the line $\theta = \pi/2$. The boundary conditions for this case are $a_l^\pm|_{\theta=\pi/2} = i^{l-1}(2l+1)/l(l+1)$ and a_l^\pm at (ρ, θ) is

$$a_l^\pm(\rho, \theta, \rho_*) = i^{l-1}(2l+1)/l(l+1)(\exp[\mp i(G[\rho_*, v] - G[\rho, v])])$$

where $\rho_* = \rho_*(\rho, \theta)$ from Eq. (5.6-3), which is Bouguer's law.

We define the functional $G[\rho, \nu]$ by

$$G[\rho, \nu] = \int_{\rho}^{\infty} \left(\frac{d \log n}{d \rho'} \right) g_l(\rho') d\rho' \quad (5.5-20)$$

For convenience in this and in the following sections we use the spectral number l and the parameter $\nu = l + 1/2$ interchangeably. The distinction between them is inconsequential because of the enormity of their values in their stationary phase neighborhoods. It is understood here that form for $g_l(\rho)$ given in Eq. (5.5-19) must be modified so that $g_l(\rho) \rightarrow 0$ for increasing $\nu > \rho$.

Using the asymptotic boundary condition for $a_l^-(\rho)$ given in Eq. (5.5-3a) and noting that $\rho \rightarrow u = kr$ asymptotically with large r , the solution $a_l^-(\rho)$ can be obtained by integrating Eq. (5.5-17), and it can be written as

$$a_l^-(\rho) = n^{1/2} i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho, \nu]) \quad (5.5-21)$$

Thus, $-G[\rho, \nu]$ is the phase retardation induced by the refractive gradient in the l th spectral component of an incoming wave, which results from travelling through a transparent, spherical symmetric, refracting medium from infinity down to a radial distance r . Initially, as $r \rightarrow \infty$, the incoming wave is planar and its spectral coefficient is given by Eq. (5.5-3a). For a homogeneous medium $G[\rho, \nu] \equiv 0$.

For thin atmospheres the term $n^{1/2}$ in Eq. (5.5-21) is essentially unity, and it will be ignored in subsequent discussions.

5.5.3 Outgoing Waves

We have a similar expression for a radial outgoing wave. In this case we let $\Delta a_l^+ = a_l^{(T)} - a_l^{(i)}$, where $a_l^{(i)}$ is the spectral coefficient of spectral number l for the outward travelling wave incident on the inner side of the boundary, and $a_l^{(T)}$ is the coefficient for the outward directed transmitted wave. The scalar potential series for both of these waves use the ξ_l^+ functions because they are outgoing waves. Also in Eq. (5.5-9b) we must change the ξ_l^- functions to ξ_l^+ functions because $a_l^{(i)}$ and $a_l^{(T)}$ are now the spectral coefficients for outgoing waves; similarly, we must change from ξ_l^+ to ξ_l^- for b_l because the reflected wave is incoming. Working through the same boundary conditions applicable to an outgoing wave and applying the same limit procedures that held for the inward case (see Eqs. (5.5-8)-(5.5-17)), one obtains a differential equation for the spectral coefficients of the outward directed wave

$$\frac{1}{a_l^+} \frac{da_l^+}{d\rho} = -i \frac{d \log n}{d\rho} g_l(\rho) \quad (5.5-22)$$

Comparing Eq. (5.5-22) with Eq. (5.5-17) (and dropping the $n^{1/2}$ term), we see that the gradients of a_l^- and a_l^+ have opposite polarities. In other words, the spatial derivative of the spectral coefficient along the radial direction of propagation is invariant to whether the wave is incoming or outgoing. This must be true from a physical consideration: the rate of phase accumulation at a given site should be the same for the radial travelling incoming and outgoing wavelets.

We see upon integrating Eq. (5.5-22) that a_l^+ will depend on the adopted value of a constant of integration. Let us fix that constant at $r = r_*$. We write a_l^+ in the form $a_l^+[\rho, \rho_*]$ to express this dual dependency; here $\rho_* = kr_*n(\rho_*)$. Integrating Eq. (5.5-22) and using Eq. (5.5-20) we obtain

$$a_l^+[\rho, \rho_*] = a_l^+[\rho_*, \rho_*] \exp[-i(G[\rho_*, \nu] - G[\rho, \nu])] \quad (5.5-23)$$

If we let $r \rightarrow \infty$, which would be appropriate when observing the refracted wave from outside the refracting medium, such as the neutral atmosphere observed from a LEO, then $G[\rho, \nu] \rightarrow 0$ and one obtains

$$a_l^+[\infty, \rho_*] = a_l^+[\rho_*, \rho_*] \exp(-iG[\rho_*, \nu]) \quad (5.5-24)$$

The phase retardation incurred by the ν th wavelet in traveling outward from r_* to infinity is $-G[\rho, \nu]$, which is the same retardation incurred by the inward traveling wavelet from infinity down to r_* .

The actual value(s) of $a_l^+[\rho_*, \rho_*]$ will depend in part on the physical properties assumed for the refracting and perhaps scattering atmosphere, and also in part on the impact parameter(s) associated with the ray(s). For example, if $dn/dr \equiv 0$ for $r < r_o$, then Eqs. (5.5-17) and (5.5-22) show that both a_l^- and a_l^+ will be constant in that region. They also must be equal there to avoid the Hankel function singularity at the origin. (Recall that the definition of the spherical Bessel function of the first kind, $\psi_l = (\xi_l^+ + \xi_l^-)/2$, which is well-behaved at the origin.) It follows in this case that $a_l^+[\rho_o, \rho_o] = a_l^-(\rho_o)$, where $a_l^-(\rho_o)$ is given from Eq. (5.5-21) and it is the applicable spectral coefficient for an incoming wave that was initially planar. At the LEO we would have in this case

$$a_l^+[\infty, \rho_o] = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho_o, v]) \quad (5.5-25)$$

Thus, $-2G[\rho_o, v]$ is the total phase delay incurred by the l th spectral coefficient of an initially plane wave with an impact parameter ρ_o as a result of the wave passing completely through an intervening medium. We will return to this topic in a later section where specific refracting and scattering models are discussed. We will also discuss later the accuracy of this particular spectral representation in terms of osculating parameters.

5.5.4 Correspondence between Cartesian and Spherical Stratified Phase Quantities

In Chapter 4 we applied the thin film concepts to a Cartesian stratified medium to solve the wave equations expressed in terms of the unitary state transition matrix $\mathbf{M}[x_2, x_1]$. Central quantities in that presentation, which are given in Eqs. (4.4-13), are the phase accumulation $\mathcal{A}(x, x_o)$ and its rate $\varpi(x)$ (with $\mu \equiv 1$) that results from the profile $n(x)$ in that Cartesian stratified medium. These are

$$\left. \begin{aligned} \mathcal{A}(x, x_o) &= k \int_{x_o}^x \varpi(x') dx', \\ \varpi(x) &= \left(n^2(x) - n_o^2 \right)^{1/2}, \quad n_o = n(x_o) \end{aligned} \right\} \quad (5.5-26)$$

Note that $\mathcal{A}(x, x_o)$ provides the total phase accumulation of the wave along the x -axis, perpendicular to the plane of stratification, from the turning point at x_o up to the altitude at x . It is an implicit function of the refractivity profile and the “angle of incidence” φ of the wave through the value of n_o , which is a constant for a particular wave (generalized Snell’s law, $n_o = n \sin \varphi$), analogous to the impact parameter ρ_* for the spherical geometry. Thus, both $\mathcal{A}(x, x_o)$ and $\varpi(x)$ depend on the angle of incidence of the wave. Defining $\rho = kx n(x)$ for the Cartesian-stratified case, it follows from Eq. (5.5-26) that \mathcal{A} may be rewritten in the form

$$\mathcal{A}(\rho, \rho_o) = \int_{\rho_o}^{\rho} \frac{\varpi}{n} d\rho' - \int_{\rho_o}^{\rho} \frac{d \log n}{d\rho'} \frac{\varpi \rho'}{n} d\rho' \quad (5.5-27)$$

The first integral provides the “geometric” phase delay ($\varpi/n = \cos \varphi$), and the second integral provides the additional phase delay resulting from the gradient of the refractivity over the interval ρ_o to ρ . The correspondence between the

spectral quantities derived in this section for spherical stratification, $g_l(\rho)$ and $G[\rho, \nu]$, and their counterparts in Cartesian stratification should be clear. It is

$$\left. \begin{aligned} G[\rho_o, \nu] - G[\rho, \nu] &\Leftrightarrow \mathcal{A}(\rho, \rho_o) - \int_{\rho_o}^{\rho} \frac{\overline{\omega}}{n} d\rho' \\ g_l(\rho) &\Leftrightarrow \frac{\rho \overline{\omega}(\rho, \rho_o)}{n(\rho)} \Rightarrow \rho \cos \varphi, \quad \sin^{-1}\left(\frac{\nu}{\rho}\right) \Leftrightarrow \varphi \end{aligned} \right\} \quad (5.5-28)$$

Note that this correspondence applies only for $\rho > \nu$. Here the angle of incidence φ in the Cartesian frame is related to the spectral number l in the spherical frame through the relation given in Eq. (4.12-8).

5.5.5 Absorption

The modified Mie scattering approach used here lends itself easily to a medium with mild absorption. Here the index of refraction has the form $\hat{n} = n(1 + i\kappa)$, where $n(r)$ is the real component and $n\kappa$ is the imaginary component. κ is the extinction coefficient and it is real. Because the refracting sphere is so large, κ must be a very small quantity or else the penetrating waves will be completely damped before escaping from the sphere. In any case, it follows from Eq. (5.5-17) that when $\kappa \neq 0$, $a_l^-(\rho)$ will have an exponentially damping component in addition to a phase delay. In this case the constant E_o , which is the amplitude of the incident wave, must be treated more carefully to account for the actual absorption through the medium. Also, in the case where the emitting GPS satellite is located at a finite distance away, ρ_{LG} , then E_o must account for the space loss in amplitude that the spherical wave emitted from the GPS satellite incurs in travelling to the LEO.

5.6 More Geometric Optics: Cumulative Bending Angle, Bouguer's Law and Defocusing

We need a few more concepts from geometric optics for incoming and outgoing waves to interpret these wave theory results using the stationary phase technique. Appendix A briefly discusses deriving the ray path in geometric optics from Fermat's principle and the Calculus of Variations. We know that the path integral for the phase delay along the ray from the observed GPS satellite to the LEO, $\int n ds$, is stationary with respect to the path followed by the signal. That is, the actual path provides a stationary value for the phase delay compared with the phase delay that would be obtained by following any neighboring path with the same end points. Here s is path length. If one applies the Calculus of Variations to this phase delay integral, then one obtains Euler's equation, which is a second order differential equation. This equation provides a necessary condition that the path must satisfy to yield a stationary value for the phase delay path integral. When the path integral is expressed in polar coordinates with r as the independent variable, then $ds = (1 + r^2 \theta'^2)^{1/2} dr$, and Euler's equation becomes

$$\frac{d}{dr} \left(\frac{\partial}{\partial \theta'} (n \sqrt{1 + r^2 \theta'^2}) \right) - \frac{\partial}{\partial \theta} (n \sqrt{1 + r^2 \theta'^2}) = 0 \quad (5.6-1a)$$

Provided that n is a function only of r , this equation may be integrated once to obtain a constant of integration

$$\frac{nr^2 \theta'}{\sqrt{1 + r^2 \theta'^2}} = n_* r_*, \quad \theta' = \frac{d\theta}{dr}, \quad n_* = n(r_*), \quad \frac{r \theta'}{\sqrt{1 + r^2 \theta'^2}} = \sin \gamma \quad (5.6-1b)$$

Here θ' is related to the slope of the ray at the point (r, θ) , and γ is the angle between the radius vector and the tangent vector of the ray. For planar approaching waves, i.e., for the case where the occulted GPS satellite is set infinitely far away in the $\theta = \pi$ direction, then $\gamma = \tilde{\alpha} + \theta$. Here $\tilde{\alpha}$ is the cumulative bending angle up to the point (r, θ) incurred by the ray relative to its original direction ($\theta = 0$) as an approaching planar wave. In Appendix A, Figure A-1, a positive value for $\tilde{\alpha}$ corresponds to a clockwise rotation of the tangent vector of the ray relative to the line $\theta = 0$. Along a ray path satisfying Euler's equation the impact parameter $\rho_* = kn_* r_*$ must be constant when n is not a function of θ . From geometric optics the differential bending angle $d\tilde{\alpha}$ over an infinitesimal length ds along the ray path expressed in polar coordinates is given by $d\tilde{\alpha} = (dn/dr) \sin \gamma ds$. Upon applying Eq. (5.6-1b) and integrating

$d\tilde{\alpha}$ along the ray path from the GPS satellite (assumed to be at infinity) to an approaching point (r, θ) one obtains

$$\left. \begin{aligned} \tilde{\alpha}[r, \theta] = \tilde{\alpha}(\rho, \rho_*) &= -\rho_* \int_{\rho}^{\infty} \frac{d \log n}{d\rho'} \frac{d\rho'}{\sqrt{\rho'^2 - \rho_*^2}} \\ \rho &= krn(r), \quad \theta_* < \theta < \pi \end{aligned} \right\} \quad (5.6-2)$$

Here θ_* marks the angular coordinate of the tangency or turning point for a particular ray with an impact parameter value of ρ_* . See Figure A-1. We note that $\theta_* = \theta_*(\rho_*)$, and also $\tilde{\alpha}$ at any approaching point (r, θ) may be considered to be a function of $\rho = rn(r)$ and the impact parameter ρ_* for the ray passing through that point. Thus, $\tilde{\alpha} = \tilde{\alpha}(\rho, \rho_*)$ and also $\tilde{\alpha} = \tilde{\alpha}[r, \theta]$. From Eqs. (5.6-1) and (5.6-2) it follows that the impact parameter $\rho_* = kr_*n(r_*)$ is given in terms of (r, θ) and $\tilde{\alpha}$ by

$$\left. \begin{aligned} \text{Incoming: } \pi > \theta \geq \theta_*: \quad \rho_* &= \rho \sin[\theta + \tilde{\alpha}(\rho, \rho_*)] \\ \text{Outgoing: } \theta_* \geq \theta > 0: \quad \rho_* &= \rho \sin[\theta + 2\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho, \rho_*)] \\ \theta_* &= \frac{\pi}{2} - \tilde{\alpha}(\rho_*, \rho_*); \quad \rho = krn(r) \end{aligned} \right\} \quad (5.6-3)$$

This is a version of Bouguer's law, which has been expressed for both an incoming ray ($\pi > \theta \geq \theta_*$) and for an outgoing ray ($\theta_* \geq \theta > 0$). By symmetry the bending angle for an outgoing ray is $2\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho, \rho_*)$, where $\tilde{\alpha}(\rho_*, \rho_*)$ is the cumulative bending up to the turning point. The constant value of ρ_* along a given ray path is the geometric optics analogue of the conservation of angular momentum in a classical mechanical system with spherical symmetry in its force field. One can solve Eqs. (5.6-2) and (5.6-3) simultaneously to obtain the values of both ρ_* and $\tilde{\alpha}(\rho, \rho_*)$ for a given position (r, θ) (which may or may not be uniquely determined, depending on the profile of $n(r)$).

We denote the coordinates (with the radial coordinate in phase units) of the LEO by (ρ_L, θ_L) . When the LEO is outside of the atmosphere where $n \equiv 1$, $\tilde{\alpha}(\rho, \rho_*) \rightarrow 0$ as $\rho \rightarrow \infty$. Therefore, Bouguer's law becomes

$$\rho_L \sin(\theta_L + \alpha_L) = \rho_* = \text{constant}, \quad \alpha_L = 2\tilde{\alpha}(\rho_*, \rho_*) \quad (5.6-4)$$

Here $\alpha_L(\rho_*)$ is the total refractive bending angle observed by the LEO. Thus, for a given LEO position (ρ_L, θ_L) , there is a one-to-one correspondence (when

spherical symmetry applies) between α_L and the impact parameter $\rho_* = kr_* n_*$. These relations are given by

$$\left. \begin{aligned} \rho_* &= \rho_L \sin(\theta_L + \alpha_L(\rho_*)), \\ \alpha_L(\rho_*) &= 2\tilde{\alpha}[\rho_*, \theta_*] = 2\tilde{\alpha}(\rho_*, \rho_*) = -2\rho_* \int_{\rho_*}^{\infty} \frac{d \log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - \rho_*^2}} \end{aligned} \right\} \quad (5.6-5)$$

where (r_*, θ_*) is the turning point for a ray with an impact parameter value of ρ_* . Given a position (ρ_L, θ_L) for the LEO, Eqs. (5.6-4) and (5.6-5) yield the values ρ_* and $\alpha_L(\rho_*)$ that must follow in order for the LEO to observe the ray (or rays if ρ_* and α_L are not uniquely determined at that position) from the GPS satellite that has been deflected as a result of refractive bending. The point of tangency of the ray on the boundary is located at an angular position $\theta = \pi/2 - \alpha_L/2$ (for an infinitely distant GPS) and at a radial position of $r = r_*$.

Figure 5-3 shows an example of the solution to Eq. (5.6-5) for an exponential refractivity profile. In this case the bending angle decreases monotonically with increasing impact parameter so that the solution is unique. If the refractivity profile causes the bending angle to exhibit reversals in slope, then there may be multiple solutions for a certain range of impact parameter values. We will discuss the question of uniqueness of the impact parameter and total bending angle for a given LEO position (ρ_L, θ_L) in Section 5.12, which also addresses multipath and ray caustics.

5.6.1 Defocusing

Because defocusing will arise when we apply the stationary phase technique in wave theory, we review it here in somewhat more detail than given in Appendix A. Defocusing arises because of dispersive refraction. Incoming rays, which were collimated prior to encountering the atmosphere, are dispersed or spread out after entering the atmosphere because of the refractive gradient. To obtain a measure of the defocusing at a given point (r, θ) , we compute the ratio of the signal flux density of an incoming wave prior to encountering the atmosphere with the signal flux density at the point (r, θ) . This ratio is readily obtained from Bouguer's law and by invoking the principle of conservation of energy. Let $\Delta\sigma$ be the perpendicular displacement in phase units between two rays (Figure A-1) in the neighborhood of the point (r, θ) that results from changing the impact parameter by an amount $\Delta\rho_*$. Conservation of energy requires (assuming complete transparency) that the power through a cross section of width $\Delta\rho_*$ in the collimated beam prior to atmospheric entry must equal the power through the cross section of width $\Delta\sigma$ at the point (r, θ) after

atmospheric entry. Thus, the ratio $\zeta = \Delta\rho_* / \Delta\sigma$ gives us the defocusing, and its square root gives the ratio of the signal amplitudes: the amplitude of the wave at (r, θ) divided by the amplitude of the collimated wave prior to atmospheric entry.

Upon differentiating Bouguer's law given in Eq. (5.6-3) with respect to ρ_* , and defining $\gamma = \theta + \tilde{\alpha}$, one obtains

$$\left. \begin{aligned} \text{Incoming: } \pi > \gamma \geq \pi/2: \Delta\rho_* = \\ & \Delta r(n + \rho n' / n) \sin \gamma + (\Delta\theta + \Delta\tilde{\alpha}(\rho, \rho_*)) \cos \gamma, \\ \text{Outgoing: } 0 < \gamma \leq \pi/2: \Delta\rho_* = \\ & \Delta r(n + \rho n' / n) \sin \gamma + (\Delta\theta + 2\Delta\tilde{\alpha}(\rho_*, \rho_*) - \Delta\tilde{\alpha}(\rho, \rho_*)) \rho \cos \gamma \end{aligned} \right\} \quad (5.6-6)$$

Here

$$\left. \begin{aligned} \Delta\tilde{\alpha}(\rho, \rho_*) &= \partial\tilde{\alpha}/\partial\rho_* \Delta\rho_* + \partial\tilde{\alpha}/\partial\rho \Delta\rho, \\ \Delta\tilde{\alpha}(\rho_*, \rho_*) &= d\tilde{\alpha}/d\rho_* \Delta\rho_*, \\ \Delta\rho &= k(n + rn') \Delta r, \quad n' = \frac{dn}{dr}, \quad k \frac{dn}{d\rho} = \frac{n'}{n + rn'}, \quad \gamma = \theta + \tilde{\alpha} \end{aligned} \right\} \quad (5.6-7)$$

From Eq. (5.6-2) it follows that

$$\left. \begin{aligned} \frac{\partial\tilde{\alpha}}{\partial\rho_*} &= - \int_{\rho}^{\infty} \frac{\rho'}{n} \frac{dn}{d\rho'} \frac{\rho' d\rho'}{(\rho'^2 - \rho_*^2)^{3/2}} = \\ & - \frac{\rho}{n} \frac{dn}{d\rho} \frac{1}{(\rho^2 - \rho_*^2)^{1/2}} - \int_{\rho}^{\infty} \frac{d}{d\rho'} \left(\frac{\rho'}{n} \frac{dn}{d\rho'} \right) \frac{d\rho'}{(\rho'^2 - \rho_*^2)^{1/2}} \end{aligned} \right\} \quad (5.6-8)$$

and

$$\frac{\partial\tilde{\alpha}}{\partial\rho} = \frac{\rho_*}{n} \frac{dn}{d\rho} \frac{1}{(\rho^2 - \rho_*^2)^{1/2}} \quad (5.6-9)$$

On the lower line of Eq. (5.6-8) the integral is well behaved and bounded as $\rho_* \rightarrow \rho$. For an exponential refractivity profile this integral is very closely equal to $-\tilde{\alpha}(\rho, \rho_*)/H$. But Eqs. (5.6-8) and (5.6-9) show that $\partial\tilde{\alpha}/\partial\rho_*$ and $\partial\tilde{\alpha}/\partial\rho$ have a $(\rho^2 - \rho_*^2)^{-1/2}$ singularity at $\rho = \rho_*$. In the defocusing expression these partial derivatives are multiplied by D , which is given by

$$D = \rho \cos \gamma = \pm (\rho^2 - \rho_*^2)^{1/2} \quad (5.6-10)$$

D is essentially the optical distance between the point (r, θ) and the tangency point on the circle of radius ρ_* . Here the plus sign is used for an outgoing ray and the minus sign for incoming ray. Multiplying the expressions in Eqs. (5.6-8) or (5.6-9) by D removes the singularity.

In Eq. (5.6-6) Δr and $r\Delta\theta$ are constrained to follow a displacement direction at (r, θ) that is perpendicular to the ray. From Figure A-1 it follows that

$$k\Delta r = \Delta\sigma \sin \gamma, \quad kr\Delta\theta = \Delta\sigma \cos \gamma \quad (5.6-11)$$

If these quantities in Eqs. (5.6-8)–(5.6-11) are substituted into Eq. (5.6-6) and Bouguer's law is used, one can show that the defocusing ratios are given by

$$\left. \begin{aligned} &\text{Incoming: } \pi > \gamma \geq \frac{\pi}{2}: \\ &\quad n \frac{d\sigma}{d\rho_*} = 1 - \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \rho \cos \gamma \\ &\text{Outgoing: } 0 < \gamma \leq \frac{\pi}{2}: \\ &\quad n \frac{d\sigma}{d\rho_*} = 1 - \left(2 \frac{d\tilde{\alpha}(\rho_*, \rho_*)}{d\rho_*} - \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \right) \rho \cos \gamma \end{aligned} \right\} \quad (5.6-12)$$

For a point located at the LEO outside of the atmosphere, $\tilde{\alpha}(\rho, \rho_*) \rightarrow 0$ as $\rho \rightarrow \infty$, and the outgoing form in Eq. (5.6-12) becomes

$$\frac{d\sigma}{d\rho_*} = 1 - 2D \frac{d\tilde{\alpha}}{d\rho_*}, \quad \tilde{\alpha} = \tilde{\alpha}(\rho_*, \rho_*) \quad (5.6-13)$$

Here D becomes the distance (in phase units) of the LEO from the Earth's limb minus $\rho_* \alpha_L(\rho_*)$. In practice the GPS satellites are not infinitely distant; their orbit radius is only about 4 Earth radii. To compensate for the wavefront curvature resulting from this finite distance, it is customary to use the “reduced distance” in Eq. (5.6-13), which is defined in the same way as the “reduced mass” in two-body dynamical systems. This is given by

$$D^{-1} = D_L^{-1} + D_G^{-1} \quad (5.6-14)$$

This definition for D follows directly from the Fresnel approximation in the thin phase screen theory (see Section 5.11). This form for D gives a slightly

more accurate measure of defocusing. It also is very useful in thin phase screen analyses for diffraction and multipath.

For a circular LEO orbit it follows from Figure A-1 that $\Delta\sigma = D\Delta\theta_L$, where $\Delta\theta_L$ is the displacement in orbital position of the LEO required to intercept two nearby rays separated in impact parameter by $\Delta\rho_*$. In this case of a circular orbit for the LEO the defocusing equation in Eq. (5.6-13) can be written as

$$\zeta^{-1} = \frac{d\sigma}{d\rho_*} = D \frac{d\theta_L}{d\rho_*} = 1 - 2D \frac{d\tilde{\alpha}}{d\rho_*} \quad (5.6-13')$$

From Eq.(5.6-2) it follows that

$$\left. \begin{aligned} \frac{d\tilde{\alpha}(\rho_*, \rho_*)}{d\rho_*} &= \frac{\tilde{\alpha}}{\rho_*} - \rho_*^2 \int_{\rho_*}^{\infty} \frac{d}{d\rho} \left(\frac{1}{\rho} \frac{d \log n}{d\rho} \right) \frac{d\rho}{\sqrt{\rho^2 - \rho_*^2}} \\ &\doteq -\rho_* \int_{\rho_*}^{\infty} \frac{1}{n} \frac{d^2 n}{d\rho^2} \frac{d\rho}{(\rho^2 - \rho_*^2)^{1/2}}, \quad \rho_* \gg 0 \end{aligned} \right\} \quad (5.6-15)$$

For an exponential refractivity profile $d^2 n / d\rho^2 = -H^{-1} dn / d\rho$, and it follows that $d\tilde{\alpha} / d\rho_* \doteq -\tilde{\alpha} / H$. In this case the defocusing is related to the deflection $2\tilde{\alpha}D$ at the LEO expressed as the number of scale heights $2\tilde{\alpha}D / H$ that refractive bending induces.

At a turning point $\gamma = \theta + \tilde{\alpha} \rightarrow \pi / 2$. One could naïvely conclude that the defocusing factor would reduce to $1 / (n_* + r_* n'_*)$, which is the reciprocal of the derivative of the impact parameter ρ_* with respect the radial coordinate kr at the turning point (r_*, θ_*) . Eqs. (5.6-8) and (5.6-10) show that the singularity in $\partial\tilde{\alpha}(\rho, \rho_*) / \partial\rho_*$ as $\rho \rightarrow \rho_*$ yields a finite contribution to the defocusing ratio at a turning point. From Eqs. (5.6-8) and (5.6-10), Eq. (5.6-12) becomes at a turning point

$$n_* \frac{d\sigma}{d\rho_*} = 1 - \frac{\rho_*}{n_*} \frac{dn}{d\rho} \bigg|_{\rho_*} = \frac{n_*}{n_* + r_* n'_*} \quad (5.6-16)$$

thus confirming our intuition.

5.7 More Asymptotic Forms

We will need asymptotic forms for the functions $g_l(\rho)$ and $G[\rho, \nu]$, defined in Eqs. (5.5-19) and (5.5-20). We also must complete the unfinished business of fixing that form for $g_l(\rho)$. The derivation obtained in Section 5.5

based on modified Mie scattering fails when $\rho < \rho^\dagger(v)$. We know from Chapter 3 that the principal contributions to the scattering integrals come from spectral coefficients with wavenumber values in the near vicinity of $l = \rho_*$. Here asymptotic forms that exploit the relatively small value of $|l - \rho_*|/\rho_*$ but the large value of ρ_* are appropriate. Therefore, we use the asymptotic forms for the spherical Hankel functions in terms of the Airy functions of the first and second kind that have been given in Eq. (5.4-1).

It then follows upon replacing the spherical Hankel functions with their Airy function asymptotic forms that for $g_l(\rho)$ one obtains from Eq. (5.5-19b)

$$g_l(\rho) \rightarrow g(\hat{y}) = \left\{ \pi K_v^2 (\text{Ai}'[\hat{y}]^2 + \text{Bi}'[\hat{y}]^2) - \hat{y} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) \left(1 - \frac{\hat{y}}{5K_v^2} + \dots \right) \right\} \quad (5.7-1)$$

Here $v = l + 1/2$. The quantities \hat{y} and K_v are defined in Eq. (5.4-3) in terms of v and ρ . When $v \approx \rho$ we can drop the \hat{y}/K_v^2 term in Eq. (5.7-1) because of the enormity of ρ_* .

It follows from Eq. (5.5-20) that the applicable asymptotic form for $G[\rho, v]$ is given by

$$G[\rho, v] \doteq \int_{\rho}^{\infty} \left(\frac{d \log n}{d \rho'} \right) g(\hat{y}') d \rho', \quad \hat{y}' = \zeta(v / \rho'), \quad \rho > \rho^\dagger(v) \quad (5.7-2)$$

Here \hat{y}' and ρ' are integration variables and they are connected through the expression for \hat{y} given in Eq. (5.4-3).

If one uses the asymptotic forms applicable to negative arguments for the Airy functions given by Eq. (3.8-7), it follows from Eq. (5.7-1) that

$$g(\hat{y}) = (\rho^2 - v^2)^{1/2} \left(1 - 1/(32\hat{y}^3) + \dots \right), \quad v < \rho \quad (5.7-3)$$

Similarly, from Eq. (5.7-2) it follows for negative arguments that

$$G[\rho, v] \doteq \int_{\rho}^{\infty} \frac{d \log n}{d \rho'} \sqrt{\rho'^2 - v^2} d \rho', \quad v < \rho \quad (5.7-4)$$

Even at $v = \rho$ this asymptotic form for $G[\rho, v]$ is very accurate. The difference between the values of $G[\rho, v]$ from Eqs. (5.7-2) and (5.7-4) is roughly $0.2\beta/(1-\beta)$, yielding a relative accuracy in most conditions of a few parts in

10^5 . A super-refractive region requires special treatment; see Sections 5.8 and 6.4.

From the cumulative bending angle given by Eq. (5.6-2), we have upon integrating on ρ_*

$$\left. \begin{aligned} \int_v^\rho \tilde{\alpha}(\rho, v') dv' &= \int_\rho^\infty \frac{d \log n}{d \rho'} \left[\sqrt{\rho'^2 - \rho^2} - \sqrt{\rho'^2 - v^2} \right] d \rho' \\ &= - \int_\rho^\infty \tilde{\alpha}(\rho', \rho') d \rho' - \int_\rho^\infty \frac{d \log n}{d \rho'} \sqrt{\rho'^2 - v^2} d \rho' \end{aligned} \right\} \quad (5.7-5)$$

From Eqs. (5.7-4) and (5.7-5) it follows that

$$G[\rho, v] \doteq - \int_v^\rho \tilde{\alpha}(\rho, \omega) d\omega - \int_\rho^\infty \tilde{\alpha}(\omega, \omega) d\omega, \quad v < \rho \quad (5.7-6)$$

We recall that $-G[\rho, v]$ is the phase delay induced by the spherical symmetric atmosphere upon a radial wavelet of wavenumber v at the radial position $\rho = krn$. Eq. (5.7-6) provides a physical interpretation for these integrals of the bending angle. This is similar in form to the stationary phase condition that we found for the thin screen model in Chapter 2. We note from Eq. (5.7-6) that $-2G[\rho, \rho]$ corresponds to the thin screen phase profile $\phi(h)$ given in Eq. (2.5-1).

Differentiating $G[\rho, v]$ given in Eq. (5.7-2) with respect to v and using the defining differential equation for the Airy functions ($\text{Ai}'' = \hat{y}\text{Ai}$, $\text{Bi}'' = \hat{y}\text{Bi}$) (and the near-linear relationship between \hat{y} and v given in Eq. (5.4-3)), one obtains

$$\frac{\partial G}{\partial v} = \int_\rho^\infty \left(\frac{d \log n}{d \rho'} \right) \frac{dg}{d \hat{y}} \frac{\partial \hat{y}}{\partial v} d \rho' \doteq -\pi K_v \int_\rho^\infty \left(\frac{d \log n}{d \rho'} \right) (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d \rho' \quad (5.7-7)$$

For negative arguments of \hat{y} we can use the negative asymptotic forms for the Airy functions; Eq. (5.7-7) becomes (alternately, one can take the partial derivative of Eq. (5.7-4))

$$\frac{\partial G}{\partial v} \doteq -v \int_\rho^\infty \frac{d \log n}{d \rho'} \frac{1}{\sqrt{\rho'^2 - v^2}} d \rho' = \tilde{\alpha}(\rho, v), \quad v < \rho \quad (5.7-8)$$

Comparing Eqs. (5.6-2) and (5.7-8), we conclude that for negative values of \hat{y} , that is, for $v < \rho$, $\partial G[\rho, v] / \partial v$ may be interpreted as *the cumulative bending angle $\tilde{\alpha}(\rho, v)$ of an incoming ray at the radial position r ($\rho = krn$) and with an impact parameter value v* . This high accuracy deteriorates only when v lies

in the immediate vicinity of ρ , that is, near a turning point (r_*, θ_*) where $\theta_* = \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$.

Figure 5-4 shows the level of this agreement in the vicinity of a turning point for an exponential refractivity profile that decreases with altitude. It roughly corresponds to the dry air refractivity profile for the Earth near sea level. Only for spectral numbers in the range $\nu > \rho - \sim 2K_\nu$ does the agreement deteriorate. Note that $\tilde{\alpha}(\rho, \nu)$ is not defined for $\nu > \rho$ ($\hat{y} > 0$) and that $\partial \tilde{\alpha}(\rho, \nu) / \partial \nu \rightarrow \infty$ as $\nu \rightarrow \rho^-$, whereas $G[\rho, \nu]$ and its derivatives are well-behaved in this neighborhood. However, the form for $G[\rho, \nu]$ derived in Section 5.5 and its asymptotic form given in Eqs. (5.7-1) and (5.7-2) changes rapidly with increasing $\nu > \rho$ because of the behavior of the Airy function of the second kind. This form fails for $\nu > \rho$. In fact, we show later that $g(\hat{y}) \rightarrow 0$ for increasing $\hat{y} > 0$, and that $G[\rho, \nu]$ approaches a constant value. The small discrepancies near the turning point result from the deviations of the Airy functions from their asymptotic forms for negative arguments, and also from the breakdown in accuracy of the spectral coefficients near a turning point when they are derived from Eq. (5.5-21).

That there should be this very close although not perfect agreement between ray quantities and spectral coefficients from wave theory when the latter are evaluated at their stationary phase values should not be too surprising. In wave theory the stationary phase process, which is discussed later, is effected over spectral number. The value $\nu = \rho_*$, which we will show to be very close to a stationary phase point in wave theory, also provides an equivalent ray in geometric optics between the GPS satellite and the point (r, θ) and with an impact parameter value of ρ_* . The ray path from geometric optics is in fact a path of stationary phase. Using geometric optics we may vary the impact parameter over impact parameter space for a ray with constrained end points. The path has a stationary value of the phase delay when the impact parameter takes on the value $\rho_* = \nu$. Any other path with the same end points in the neighborhood of the actual ray path would present the observer at (r, θ) with a phase delay that differed from the observed phase delay by an amount that has only a second order dependency on the coordinate and slope deviations of the alternate path. This second order variation would be due to a deviation in the value of the impact parameter. We will show that when stationary values are assumed by the spectral number in wave theory and by the impact parameter in geometric optics, a close correspondence results.

Consider next the variability of $G[\rho, \nu_o]$ with r for a fixed value of the wavenumber ν_o . At a point (r, θ) on the approaching side at large distances where $n \rightarrow 1$ we set $kr \sin \theta = \nu_o$; that is, ν_o becomes the impact parameter for the ray passing through the point (r, θ) . Figure 5-6 shows an example of

$G[\rho, \nu_o]$ versus ρ in the vicinity of $\rho = \nu_o$. Here an exponential refractivity profile decreasing with altitude is used, but the altitude range shown in the figure is very narrow, $\sim 7K_{\nu_o}/k$, or about 100 meters for the parameter values shown. This is a small fraction of the refractivity scale height H (7 km) adopted in the figure. On the left side of the figure where $\hat{y} < 0$, that is, where $\rho > \nu_o$, the $g(\hat{y})$ function is slowly decreasing with increasing \hat{y} (decreasing ρ) (see Figure 5-2), and $G[\rho, \nu_o]$ decreases with increasing \hat{y} because the gradient of the refractivity is negative. On the right side where $\hat{y} > 0$, $g(\hat{y})$ breaks sharply negative because of the rapid growth of the Airy function of the second kind in this region. Therefore, $G[\rho, \nu]$, according to the (erroneous) form given in Eq. (5.7-2), rapidly increases with increasing \hat{y} because the refractivity gradient is negative in the example shown in Figure 5-6. Therefore, $G[\rho, \nu]$ has a stationary value at the zero crossing of $g(\hat{y})$ near $\rho = \nu_o$. The correct form for $G[\rho, \nu]$ approaches a constant for increasing $\hat{y} > 0$. This point where $g(\hat{y}) = 0$ marks a stationary phase point for $a_l^-(\rho)$ (with $l = \nu_o + 1/2$). It also is a stationary point for $a_l^+(\rho)$. Later, we will set $a_l^+ = a_l^-$ at this point to ensure no singularity at the origin from the spherical Hankel function.

For other refractivity profiles there may be other stationary points for $G[\rho, \nu]$, but their occurrence and location are dependent on the functional form of the index of refraction. However, there is always one near $\rho = \nu_o$ (unless $dn/d\rho \equiv 0$), reflecting the “deepest penetration” by the corresponding ray into the sphere. This region ($\hat{y} > 0$) corresponds to refractivity features lying *below* the impact parameter, or the point of closest approach of the associated ray. Such features lying below the point of closest approach are not “felt” by the ray. Refractivity features lying near or above the point of closest approach ($\hat{y} \leq 0$) are “felt” twice by the ray, incoming and outgoing; these can have a prominent effect on the shape of $G[\rho, \nu]$ depending on the actual refractivity profile.

From the definition of $G[\rho, \nu]$ in Eq. (5.7-2) and the asymptotic form for $g_l(\rho) \rightarrow g(\hat{y})$ in Eq. (5.7-1), it follows that

$$\left. \begin{aligned} \frac{\partial G}{\partial \rho} &= -\frac{d \log n}{d\rho} g(\hat{y}) \doteq \\ &- \pi K_v^2 \frac{d \log n}{d\rho} \left(\text{Ai}'[\hat{y}]^2 + \text{Bi}'[\hat{y}]^2 - \hat{y} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) \right) \end{aligned} \right\} \quad (5.7-9)$$

Upon setting $\partial G / \partial \rho = 0$, it follows that either $g(\hat{y}) = 0$ or $dn/d\rho = 0$, or both are zero. Consider first $g(\hat{y}) = 0$ in Eq. (5.7-9). This yields a stationary point for \hat{y} that is given by

$$\hat{y}^\dagger = \frac{\text{Ai}'[\hat{y}^\dagger]^2 + \text{Bi}'[\hat{y}^\dagger]^2}{\text{Ai}[\hat{y}^\dagger]^2 + \text{Bi}[\hat{y}^\dagger]^2} = 0.44133... \quad (5.7-10)$$

From the properties of the Airy functions it can be shown that this value of \hat{y}^\dagger is unique. We have seen this quantity before in Section 4.6 in regard to turning points for waves traveling in a Cartesian stratified Airy medium.

Upon using Eq. (5.4-3), Eq. (5.7-9) yields a stationary phase point $\rho = \rho^\dagger$ that is a function of v_o and which is given by

$$\rho^\dagger \doteq v_o - K_{\rho^\dagger} \hat{y}^\dagger, \quad K_{\rho^\dagger} = \left(\frac{\rho^\dagger}{2} \right)^{\frac{1}{3}} \quad (5.7-11)$$

Thus $|\rho^\dagger - v_o|/v_o = y^\dagger / 2K_{\rho^\dagger}^2 \sim 10^{-6}$ for $v_o \sim 10^8$.

Continuing to the second derivative $\partial^2 G / \partial \rho^2$, and evaluating it at $\rho = \rho^\dagger(v_o)$, it follows that

$$\left. \frac{\partial^2 G}{\partial \rho^2} \right|_{\rho^\dagger} \doteq -\pi K_v \frac{d \log n}{d \rho} \bigg|_{\rho^\dagger} \left(\text{Ai}[\hat{y}^\dagger]^2 + \text{Bi}[\hat{y}^\dagger]^2 \right) \quad (5.7-12)$$

Setting v_o equal to $kr_o n(r_o)$, adopting the Earth's dry air conditions at sea level and using 19 cm for the GPS wavelength, $\partial^2 G / \partial \rho^2 \big|_{\rho^\dagger} \sim 10^{-6}$ rad. It follows that G will vary from its stationary phase value by the order of 1 radian when $G|\rho - \rho^\dagger|/\rho^\dagger \sim 10^{-5}$; in other words, the stationary phase neighborhood about $\rho = \rho^\dagger$ is very narrow (a few dekameters) when ρ^\dagger is so large ($\sim 2 \times 10^8$).

In stationary phase theory to be discussed in a later section, we will use $\partial^2 G[\rho, v] / \partial v^2$, which is related to defocusing. A comparison of this second derivative with $\partial \tilde{\alpha}(\rho, v) / \partial v$, also used in geometric optics for defocusing, is shown in Figure 5-5 for the same conditions given in Figure 5-4. These also agree closely except near a turning point. Accordingly, we expect the accuracy of the amplitude predicted by the osculating parameter technique to degrade for $\hat{y} > -2$. Numerical results verify this threshold. At a turning point the correct value for the defocusing in a refracting medium without scattering is given by Eq. (5.6-16). At a turning point the stationary phase value in spectral number using the osculating parameter approach is equivalent to $\hat{y} \approx 0$; it predicts unity there for the defocusing.

5.7.1 Equating $dG[\rho^\dagger(v), v]/dv$ to $\tilde{\alpha}(v, v)$ and $d^2G[\rho^\dagger(v), v]/dv^2$ to $d\tilde{\alpha}(v, v)/dv$

We note in Figure 5-4 at the turning point that the value of $\partial G[\rho_o, v]/\partial v$ at $v = v^\dagger = \rho_o + K_{\rho_o} \hat{y}^\dagger$ is very close to the value of $\tilde{\alpha}(\rho_o, \rho_o)$. In fact, the actual numbers for Figure 5-4 are 10.114 mrad for $\tilde{\alpha}(\rho_o, \rho_o)$ and 10.115 mrad for $\partial G[\rho_o, v]/\partial v$. Since $\rho = \rho^\dagger(v)$ marks a stationary point for $G[\rho, v]$ where $\partial G/\partial \rho = 0$, it follows that

$$\frac{dG^\dagger(v)}{dv} \equiv \frac{dG[\rho^\dagger(v), v]}{dv} = \left(\frac{\partial G[\rho, v]}{\partial v} \right)_{\rho=\rho^\dagger} \quad (5.7-13)$$

Hence, $dG^\dagger(v)/dv$ also will be very close in value to $\tilde{\alpha}(v, v)$. We show in Appendix J that

$$\frac{dG^\dagger(v)}{dv} = \tilde{\alpha}(v, v) + O\left[\rho \frac{d^2 n}{d\rho^2}\right], \quad G^\dagger(v) = G[\rho^\dagger(v), v] \quad (5.7-14)$$

Therefore, the difference between $dG^\dagger(v)/dv$ and $\tilde{\alpha}(v, v)$ depends on the curvature in the refractivity profile. For the thin atmosphere conditions shown in Figure 5-4 it can be shown that $\rho_o(d^2 n/d\rho^2) \approx 10^{-3}$ mrad, which we can ignore. The difference becomes significant when near-super-refractivity conditions are encountered. See Appendix J.

Similarly, it can be shown that

$$\left. \begin{aligned} \frac{d^2 G^\dagger(v)}{dv^2} &\equiv \frac{d^2 G[\rho^\dagger(v), v]}{dv^2} = \left(-\frac{\partial^2 G[\rho, v]}{\partial \rho^2} + \frac{\partial^2 G[\rho, v]}{\partial v^2} \right)_{\rho=\rho^\dagger} \\ &= \frac{d\tilde{\alpha}(v, v)}{dv} + O\left[\frac{d^2 n}{d\rho^2}\right] \end{aligned} \right\} \quad (5.7-15)$$

For thin atmosphere conditions the curvature term here also can be dropped. Eqs. (5.7-14) and (5.7-15) will play key roles when we apply the stationary phase technique in wave theory to outgoing waves. Eq. (5.7-14) is related to the ray bending angle and Eq. (5.7-15) is related to defocusing.

Regarding the near-equivalence of $dG[\rho^\dagger(v), v]/dv$ and $\tilde{\alpha}(v, v)$, we note a property of Bauer's identity. It is given in cylindrical coordinates in Eq. (4.10-9). We can transform $\theta \rightarrow \theta + \alpha$ to obtain

$$\exp[i\rho \cos(\theta + \alpha)] = \sum_{l=-\infty}^{l=\infty} i^l J_l(\rho) \exp[il(\theta + \alpha)] \quad (5.7-16)$$

Noting that $G[\rho_o, \nu]$ shows up in the exponential term in the spectral series representations for the electromagnetic field, its variability with spectral number will be related to the angle α above. Thus, we would expect $\partial G[\rho_o, \nu] / \partial \nu$ to be closely related to an angle, which turns out to be the bending angle.

5.7.2 Fixing the Form for $g(\hat{y})$ when $\hat{y} > 0$ by Asymptotic Matching

We have noted the failure of the modified Mie scattering approach to secure the correct form for $g(\hat{y})$ when $\hat{y} \gg 0$. Eq. (5.7-1) predicts that $g(\hat{y}) \rightarrow \infty$ for increasing $\hat{y} > 0$; in fact, $g(\hat{y})$ should approach zero. We can use the form for $g(\hat{y})$ obtained from an Airy layer as guidance. In an Airy layer with a boundary at $r = r_0$ the profile for the index of refraction is given by $n^2 = n_0^2 + 2n_0 n' (r - r_0)$, where n_0 and n' are constants. The quadratic term $(n' r_0)^2 ((r - r_0) / r_0)^2$ is negligible.

In Chapter 4, Section 4.12, we showed that the solutions to the wave equations in a spherical Airy layer are given by the Airy functions. Let $U^\pm(\tilde{y})$ be a solution for the scalar field, top sign for an outgoing wave, bottom sign for an incoming wave. When $\beta < 1$ we have from Eqs. (4.10-3), (4.12-4)-(4.12-6)

$$U^\pm(\tilde{y}) = c^\pm (\text{Ai}[\tilde{y}] \mp i \text{Bi}[\tilde{y}]) \quad (5.7-17)$$

where c^\pm are complex constants obtained from matching this Airy function solution and its derivative at the boundary $r = r_0$ with the incoming and outgoing wave forms applicable on the other side of the boundary. The argument of the Airy functions is given by

$$\left. \begin{aligned} \tilde{y} &= \frac{1}{K_{\rho_0} |1 - \beta|^{2/3}} ((\beta - 1)(\tilde{\rho} - \tilde{\rho}_0) + l - \tilde{\rho}_0), \\ \beta &= -\frac{n' r_0}{n_0}, \quad \tilde{\rho} = k n_0 r \end{aligned} \right\} \quad (5.7-18)$$

These forms also apply in a super-refracting medium where $\beta > 1$. Also, it is easily shown that in an Airy layer

$$\left. \begin{aligned} \rho = kn(r) &= \rho_0 + k(n_0 + n'r_0)(r - r_0) + kn'(r - r_0)^2 \\ &\doteq \tilde{\rho}_0 + (1 - \beta)(\tilde{\rho} - \tilde{\rho}_0) - \beta(\tilde{\rho} - \tilde{\rho}_0)^2 / \tilde{\rho}_0 \end{aligned} \right\} \quad (5.7-19)$$

The quadratic term is negligible and it is dropped. Note that $\rho_0 = \tilde{\rho}_0$. It follows that the arguments of the Airy functions are related by

$$\hat{y} \doteq \frac{1}{K_{\rho_0}}(l - \rho) = \tilde{y}|1 - \beta|^{2/3} \quad (5.7-20)$$

From Eq. (5.7-17) it follows that the phase $\psi_l^{A^-}$ of the incoming Airy function solution (for $\beta < 1$) is given by

$$\psi_l^{A^-} = \tan^{-1} \left(\frac{\text{Bi}[\tilde{y}]}{\text{Ai}[\tilde{y}]} \right) + \text{constant} \quad (5.7-21)$$

Similarly, by expressing an incoming spherical Hankel function $\xi_l^-(\rho)$ in terms of its Airy function asymptotic form, its phase is given by

$$\psi_l^{H^-} = \tan^{-1} \left(\frac{\text{Bi}[\tilde{y}]}{\text{Ai}[\tilde{y}]} \right) + \text{constant} \quad (5.7-22)$$

The difference $\psi_l^{A^-} - \psi_l^{H^-}$ is the phase accumulation in the l th spectral coefficient $a_l^-(\rho)$ for an Airy layer. We denote this phase of the spectral coefficient by the function $\vartheta_l^-(\rho) = \psi_l^{A^-}(\rho) - \psi_l^{H^-}(\rho)$. Taking the derivative $\partial \vartheta_l^- / \partial \rho$, we obtain

$$\frac{\partial \vartheta_l^-}{\partial \rho} = \frac{-1}{\pi K_{\rho_0}} \left(\frac{\tilde{y}}{\text{Ai}[\tilde{y}]^2 + \text{Bi}[\tilde{y}]^2} - \frac{\hat{y}}{\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2} \right) \frac{1}{\hat{y}} \quad (5.7-23)$$

We note that $\partial \vartheta_l^- / \partial \rho$ rapidly approaches zero for increasing $\hat{y} > 0$.

We compare this expression with $\partial G[\rho, \nu] / \partial \rho$ given in Eq. (5.7-9). For negative \hat{y} we replace the Airy functions in Eq. (5.7-23) with their asymptotic forms. Then, using Eq. (5.7-20), we obtain

$$\left. \begin{aligned} \frac{\partial \vartheta_l^-}{\partial \rho} &\rightarrow \frac{-1}{K_{\rho_0}} \left((-\tilde{y})^{3/2} - (-\hat{y})^{3/2} \right) \frac{1}{\hat{y}} \\ &= \frac{\beta}{(1-\beta)K_{\rho_0}} \sqrt{-\hat{y}} \rightarrow \frac{-n'}{1-\beta} \sqrt{\rho^2 - v^2} \end{aligned} \right\} \quad (5.7-24)$$

Whereas, from Eq. (5.7-9) $\partial G[\rho, v] / \partial \rho$ becomes

$$\frac{\partial G}{\partial \rho} = -\frac{d \log n}{d \rho} g(\hat{y}) \rightarrow -2K_v^2 \frac{d \log n}{d \rho} \sqrt{-\hat{y}} = \frac{-n'}{1-\beta} \sqrt{\rho^2 - v^2} \quad (5.7-25)$$

These asymptotic forms are identical for decreasing $\hat{y} < 0$.

Eq. (5.7-23) for $\partial \vartheta_l^- / \partial \rho$ applies only to an Airy layer, but it is essentially exact and it applies without restriction on the value of \hat{y} . On the other hand, $\partial G[\rho, v] / \partial \rho$ obtained from Eq. (5.7-9) is a general form applying to any physical profile for $n(r)$, but it fails for increasing $\hat{y} > 0$. Clearly, we have a potential match made in heaven. Over the troublesome interval, say beginning at $\hat{y} > -2$, we can use the form for $\partial \vartheta_l^- / \partial \rho$ given in Eq. (5.7-23). The range $-2 \leq \hat{y} \leq 2$ corresponds to a spatial range of $4k^{-1}K_{\rho_0} \approx 60$ m. Approximating the index of refraction profile by a constant gradient n' should be fairly accurate in most circumstances, especially since the phase variability is rapidly dying out with increasing $\hat{y} > 0$.

If a single Airy layer is not sufficient, then we can form a concatenated series of Airy layers near a turning point. Any physical refractivity profile can be approximated by a series of Airy layers. In this case β and \tilde{y} would be discontinuous according to Eqs. (5.7-18) and (5.7-20) across each boundary between Airy layers (see Section 4.7). Applying the continuity conditions to the wave functions and their derivatives across each boundary ties the Airy function solutions together for the different layers, which enables one to derive the phase in any layer. For a given spectral number, whenever one reaches downward through successive Airy layers to a radial distance $\rho < \rho^\dagger(v)$, \tilde{y} will become positive there and $\partial \vartheta_l^- / \partial \rho$ for that Airy layer rapidly approaches zero for increasing $\hat{y} > 0$.

Figure 5-7 shows two comparisons of $\partial \vartheta_l^- / \partial \rho$ and $\partial G[\rho, v] / \partial \rho$ versus \hat{y} . In panel (a) the refractivity profile corresponds to dry air at sea level. But in (b) a rather hefty value of $\beta = 0.9$ has been used, hardly thin atmosphere material.

5.7.3 Behavior of the WKB Solution at a Turning Point

The phase of a wave approaches a constant value as a function of ρ below a turning point. Its behavior can be quantified by examining the WKB solutions to the radial wave equation in Eq. (5.3-5), which are given in Eq. (5.3-10) with $f_l(u) = (n^2 u^2 - l(l+1)) / u^2$. The incoming and outgoing WKB solutions, $W_l^-(u)$ and $W_l^+(u)$, can be linearly combined to yield in the $f_l < 0$ regime where $\rho < v$ an exponentially damping solution for decreasing $\rho < v$, and also an exponentially increasing solution. Using the connection formulas [6] between the WKB solutions for these two regimes, $\rho < v$ and $\rho > v$, we have

$$\left. \begin{array}{ccc} \underline{f_l(u) < 0} & \underline{f_l(u_o) = 0} & \underline{f_l(u) > 0} \\ \frac{1}{2(-f_l)^{\frac{1}{4}}} \exp\left(-\int_u^{u_o} \sqrt{-f_l} du\right) & \rightarrow & \frac{1}{(f_l)^{\frac{1}{4}}} \cos\left(\int_{u_o}^u \sqrt{f_l} du - \frac{\pi}{4}\right) \\ \frac{\sin\left(\varphi + \frac{\pi}{4}\right)}{(-f_l)^{\frac{1}{4}}} \exp\left(\int_u^{u_o} \sqrt{-f_l} du\right) & \leftarrow & \frac{1}{(f_l)^{\frac{1}{4}}} \cos\left(\int_{u_o}^u \sqrt{f_l} du + \varphi\right) \end{array} \right\} \quad (5.7-26)$$

where φ is an arbitrary phase that is not too close in value to $-\pi/4$ ¹. For $f_l = u$ and $\varphi = +\pi/4$, these are the leading terms of the asymptotic forms for the Airy functions of the first and second kind.

It follows that one solution to the wave equation is exponentially increasing for decreasing $\rho < v$, and the other is exponentially damping to zero. The asymptotic form of the incoming wave, U_l^- , in the $\rho < v$ regime is a linear combination of these two exponential solutions. Therefore, the phase of U_l^- must approach a constant value with decreasing $\rho < v$. The Airy layer analysis just discussed shows that the phase of U_l^- *rapidly* approaches a constant; therefore, $\partial G / \partial \rho$ must rapidly approach zero. Expanding f_l in powers of

¹ The arrow in each of these two connection formulas indicates the applicable direction of information transfer. For example, continuing the exponentially damping solution in the $f < 0$ regime into the $f > 0$ regime leads to a stable sinusoidal solution in that regime with “twice” the amplitude and a phase offset of $-\pi/4$. But the reverse can not be guaranteed. Inaccuracy in the numerical computation of the solution to the wave equation starting in the $f > 0$ regime and integrating downward into the $f < 0$ regime, even with the phase set equal to $-\pi/4$ exactly, leads inevitably to a numerical solution that blows up for decreasing $\rho < v$.

$u - u_o$ and setting $n_o u_o = l$, leads to $f_l = -\hat{y} / K_{\rho_o}^2 + O[\hat{y} / K_{\rho_o}^2]^2$. The exponential terms in the WKB solutions become $\exp(\pm 2\hat{y}^{3/2} / 3)$. For $\hat{y} = 2$ the ratio of their amplitudes is about 100:1; for $\hat{y} = 3$ it is 2000:1.

5.7.4 Setting $G[\rho, \nu]$ for $\rho \leq \rho^\dagger(\nu)$

In the sequel, we have taken a simpler approach, in view of Eq. (5.7-6) and also the very close agreement between $dG[\rho^\dagger(\nu), \nu] / d\nu$ and $\tilde{\alpha}(\nu, \nu)$. We let $G[\rho, \nu]$ run its course based on Eq. (5.7-2) for $\rho^\dagger(\nu) \leq \rho < \infty$. It has a stationary value at $\rho = \rho^\dagger(\nu)$. Then we set $G[\rho, \nu] = G[\rho^\dagger(\nu), \nu] = G^\dagger(\nu)$ for $\rho \leq \rho^\dagger(\nu)$, a constant value for a given value of ν . In summary, we modify Eq. (5.7-2) as follows

$$G[\rho, \nu] = \begin{cases} \int_{\rho}^{\infty} \left(\frac{d \log n}{d\rho'} \right) g(\hat{y}') d\rho', & \rho \geq \rho^\dagger(\nu) \\ \int_{\rho^\dagger}^{\infty} \left(\frac{d \log n}{d\rho'} \right) g(\hat{y}') d\rho', & \rho \leq \rho^\dagger(\nu) \end{cases} \quad (5.7-27)$$

Also, we have

$$\frac{\partial G[\rho, \nu]}{\partial \rho} = \begin{cases} - \left(\frac{d \log n}{d\rho'} \right) g(\hat{y}'), & \rho \geq \rho^\dagger(\nu) \\ 0, & \rho \leq \rho^\dagger(\nu) \end{cases} \quad (5.7-28)$$

This arrangement, which involves some approximation (see Figure 5-7), also simplifies and clarifies the strategy for matching the incoming and outgoing spectral coefficients to eliminate the Hankel function singularity at the origin. We discuss this further in Section 5.8.

5.8 Spectral Representation of an Electromagnetic Wave in a Spherical Stratified Medium

We will need to distinguish between an incoming region in the medium and an outgoing region. Figure 5-8 provides an example of the simplest topology for these regions. The boundary between these regions is given by the locus of points (r_*, θ_*) , which defines the turning point for each ray, where for a spherical symmetric medium $E_\theta(r_*, \theta_*) = 0$. In geometric optics $\theta_* = \pi / 2 - \tilde{\alpha}(\rho_*, \rho_*)$, which is obtained from Eq. (5.6-5). By “incoming” we mean the field at any point where the Poynting vector $\mathbf{S} = c(\mathbf{E} \times \mathbf{H}) / 4\pi$ for the

wave is such that $\mathbf{S} \cdot \mathbf{r} < 0$. For the simple topology shown by Figure 5-8, for example, the incoming region includes any point with a radial position r and an angular position θ that satisfies the condition $\theta_* < \theta < \pi$.

When the profile of the refractivity gradient has reversals in polarity, these regions may not be so simply connected. When a scattering surface is present, one could obtain both incoming and outgoing waves at the same point. Figure 5-9 sketches a more complicated refractivity profile that also produces coincident incoming and outgoing waves in certain regions.

We first develop the spectral coefficients for the pure refraction case and we assume that no scattering surfaces are present in the medium. Scattering or diffraction occurs where sharp changes in gradient or discontinuities in the refractivity occur. We will deal with those cases later. In particular, we assume that $n(r)$ and its derivatives are continuous throughout the medium and that the simple topology of the kind shown in Figure 5-8 applies.

Let us evaluate the electric field vector $\mathbf{E}(r, \theta)$ at a point within the refracting medium in terms of its incoming and outgoing spectral coefficients. We assume that the planar asymptotic boundary condition applies to the approaching wave and, therefore, a_l^- is obtained from Eq. (5.5-21). The proper functional form for a_l^+ is more problematical and we defer that to later in this section. From Eqs. (5.3-6), (5.5-2) and (5.5-8) we have the following spectral representation for the in-plane ($\phi = 0$) radial and transverse components of the electric field for the *TM* wave

$$\left. \begin{aligned} E_r &= \frac{E_o}{2\rho^2} \sum_{l=1}^{\infty} l(l+1) (a_l^- \xi_l^- + a_l^+ \xi_l^+) P_l^l(\cos \theta) \\ E_\theta &= \frac{E_o}{2\rho} \sum_{l=1}^{\infty} (a_l^- \xi_l'^- + a_l^+ \xi_l'^+) \frac{\partial}{\partial \theta} P_l^l(\cos \theta) \end{aligned} \right\} \quad (5.8-1a)$$

Because of the enormity of ρ and of l from where essentially all contributions to these summations originate, we can replace these summations by integrals, which are given by

$$\left. \begin{aligned} E_r &= \frac{E_o}{2\rho^2} \int_0^\infty v^2 (a_l^- \xi_l^- + a_l^+ \xi_l^+) P_l^l(\cos \theta) dv \\ E_\theta &= \frac{E_o}{2\rho} \int_0^\infty (a_l^- \xi_l'^- + a_l^+ \xi_l'^+) P_l^{l'}(\cos \theta) dv \end{aligned} \right\} \quad (5.8-1b)$$

A similar set holds for the magnetic field. In the limit as n assumes a constant value throughout the medium, $a_l^+ \rightarrow a_l^- \rightarrow a_l$, and Eq. (5.8-1) reduces to the

collimated form in Eq. (5.5-1) with $a_l = i^{l-1}(2l+1)/l(l+1)$, as it must. Also, using stationary phase arguments we will show later that in the incoming region well away from a turning point only the $a_l^-(\rho)\xi_l^-(\rho)$ term contributes significantly to the scattering integral; in the outgoing region only the $a_l^+\xi_l^+$ term contributes. Only for points very near a turning point must we include contributions from both spherical Hankel functions. This means that except near a turning point, we may replace the terms $a_l^\pm \xi_l^\pm / 2$ with $a_l^\pm \psi_l$ for computational purposes, which decays to zero rapidly with increasing $l > \rho$.

5.8.1 Behavior of $\partial G / \partial \nu$

The spectral representation for the electric field in Eq. (5.8-1b) involves integrals over spectral number. To evaluate these integrals using the stationary phase technique as an aid, we need to find those spectral neighborhoods where the phase accumulation of the integrands is stationary. In this regard, we study the variability of $G[\rho, \nu]$ with spectral number, $\partial G / \partial \nu$, in terms of the refractivity profile that determines it. We first look at $\partial G / \partial \nu$ for two different refractivity profiles. Both examples adhere to the thin atmosphere assumptions. For Case a), n is exponentially distributed so that $d\alpha/dr$ is monotonic negative with altitude; therefore, no multipath nor shadow zone situations arise. For this case $\partial G / \partial \nu$ has already been shown in Figure 5-4. In Case b), n has a Gaussian distribution; multipath, shadow zones and caustics are prominent features for this distribution. These two profiles for the index of refraction are given by

$$\left. \begin{aligned} \text{a) } n &= 1 + N_o \exp[-(\rho - \rho_o) / H], \\ \text{b) } n &= 1 + N_w \exp[-(\rho - \rho_w)^2 / 2H_w^2] \end{aligned} \right\} \quad (5.8-2)$$

Case b) is useful for study of spherical shell structures embedded in an ambient profile such as that given by Case a). Case b) could be used to describe the refractivity profile of a sporadic E layer in the ionosphere or a marine layer in the lower troposphere. Case a) can be used to describe the Earth's refractivity profile for dry air, and the values used for N_o and H in these examples correspond roughly to dry air refractivity ($N_o = 270 \times 10^{-6}$) and scale height ($k^{-1}H = 7 \text{ km}$) at sea level. For computational convenience we now have written these refractivity profiles in terms of $\rho = kr n(\rho)$, i.e., r is an implicit function of ρ^2 .

² The extraction of n given a value of $u = kr$ through iteration of Eq. (5.8-2) is cumbersome because $\rho = un(\rho)$. However, this form for an exponential distribution in

For the exponential refraction profile given by Case a), it can be shown using the thin atmosphere assumption in the defining path integral for $\tilde{\alpha}(\rho, \nu)$ given in Eq. (5.6-2) (see also Appendix A, Eq. (A-30), that $\tilde{\alpha}(\rho, \nu)$ is given by

$$\tilde{\alpha}(\rho, \nu) \doteq (n(\rho) - 1) \sqrt{\frac{\pi \rho}{2H}} \left(1 - \operatorname{erf} \left[\sqrt{\frac{\rho^2 - \nu^2}{2\rho H}} \right] \right), \quad \nu < \rho \quad (5.8-3)$$

where $n(\rho)$ is given by Eq. (5.8-2a). This expression is very accurate when compared to the value from the path integral in Eq. (5.6-2). For the decreasing exponential distribution given in Eq. (5.8-2a) the difference is less than 1% for dry air. Eq. (5.8-3) accounts for both the first and second order ray path bending effects when the refractivity profile in Eq. (5.8-2a) is used. For the dry air component of the Earth's atmosphere the first order bending term $(\sqrt{2} - 1)un'/n$ contributes less than 10% of the total. The second order ray path bending effect for dry air at sea level amounts to roughly 1% of the total.

Figure 5-10 shows the variation of $\partial G / \partial \nu$ with ν while holding ρ fixed for Case a). It was obtained from the integral in Eq. (5.7-7). For $\nu < \rho$ we have already noted from Eqs. (5.6-2) and (5.7-8), and from Figure 5-4, that $\partial G / \partial \nu \doteq \tilde{\alpha}(\rho, \nu)$, which holds very accurately for all values of ν nearly up to $\nu = \rho$. Numerical integrations of $\partial G / \partial \nu$ from Eq. (5.7-7) and $\tilde{\alpha}(\rho, \nu)$ from Eq. (5.6-2) show for both Cases a) and b) microradian level agreement when

terms of ρ has some advantages. It results in the very simple form for $\tilde{\alpha}(\rho, \nu)$ given in Eq. (5.8-3), which closely agrees with the numerical integration version for $\tilde{\alpha}(\rho, \nu)$ given in Eq. (5.6-2). Provided $|un'/n| < 1$, there is a unique relationship between u and ρ . Eq. (5.8-3) includes the effects of second order ray path bending in the path integral in Eq. (5.6-2). When a positional exponential form is used instead, $n(u) = 1 + N_o \exp[-(u - u_o)/H_u]$, the form for $\tilde{\alpha}(u, \nu)$ requires a series in powers of (un'/n) to account for higher order ray path bending. This has been discussed in Appendix A. For a given value of scale height $H_\rho = H$, the version in Eq. (5.8-2a) gives a smaller radial gradient than the positional exponential version (about 20% less for dry air at sea level). This is seen by noting that for the two versions of the exponential distributions, one obtains for dry air at sea level:

$$H_u / H_\rho \doteq (d \log n / d \log \rho) / (d \log n / d \log u) = d \log u / d \log \rho = 1 - \rho n' / n \approx 1.2.$$

One can adjust the values of the parameters N_o and H in Eq. (5.8-2a) to attain a close, but not exact, match with the profile from the positional exponential form; also, the bending angle profiles from the two versions can be matched rather closely.

For near-super-refractivity situations where $un'/n \rightarrow -1$, then the functional form $n(\rho)$ becomes inconvenient because $dn/d\rho = n'/(1 + un') \rightarrow \infty$. But the defining integral for $G[\rho, \nu]$ readily allows a change of variable to u .

the inequality $v - \rho < \sim -2K_\rho$ and the thin atmosphere assumptions are satisfied.

The three curves for $\partial G / \partial v$ in Figure 5-10 correspond to three different radial positions which are defined, respectively, by $\rho - \rho_o = 0, H/2, H$, and $\rho = un(\rho)$. The dashed curves in Figures 5-10 and 5-11 mark the value of $\partial G / \partial v$ for $v = \rho_* + \hat{y}^\dagger K_{\rho_*}$ as a function of ρ_* , which is a stationary point for $G[\rho_*, v]$. We have shown that at this stationary point $dG / dv = \partial G / \partial v \doteq \tilde{\alpha}(\rho, \rho)$ with high accuracy, where $\tilde{\alpha}(\rho, \rho)$ is given from either Eqs. (5.6-2), or for Case a) from Eq. (5.8-3) also.

Figure 5-11 shows the resulting curves for $\partial G / \partial v$ obtained from the integral in Eq. (5.7-7) when the Gaussian refractivity profile in Eq. (5.8-2b) is used. If this profile also is applied in Eq. (5.6-2) to obtain $\tilde{\alpha}(\rho, \rho)$, a bipolar refractive bending angle profile results, which is shown in Appendix E, Figure E-1, and also as the dashed curve in Figure 5-11. The bending angle profile mimics the shape of $dn / d\rho$ but it is modified to reflect the geometry of a ray transecting a spherical shell. The intersection of the $\partial G / \partial v$ curve with the $\tilde{\alpha}(\rho, \rho)$ curve in Figure 5-11 also occurs very near the point $v = \rho + \hat{y}^\dagger K_\rho$. The point where the polarity change occurs for $\partial G / \partial v$ at its initial break point depends on the location of the center of the Gaussian distribution relative to ρ_o . For $\rho_w < \rho_o$ the initial break is negatively directed, but for $\rho < \rho_w$ the slope of $n(\rho)$ becomes negative and $\partial G / \partial v$ will then break positive. Physically, this regime where the variability of $G[\rho, v]$ approaches zero, see Eqs. (5.7-26)-(5.7-28), corresponds to the Gaussian layer being located below the level, $\rho_o = v$; the layer can not be “sensed” by a ray with an impact parameter $v > \rho_o$. We will show later that there are no stationary phase points in spectral number for this regime.

5.8.2 Accuracy of the Osculating Parameter Technique

To check the accuracy of the spectral representation used in this section, as given in Eq. (5.8-1), we again use the Airy layer model for a refracting medium with spherical stratification. We embed this layer in an otherwise homogeneous medium. These analytic solutions can be compared with the osculating parameter and numerical solutions. We let

$$\left. \begin{aligned} n^2 &= n_A^2 + 2n_A \frac{dn}{d\tilde{\rho}}(\tilde{\rho} - \rho_A), \quad \frac{dn}{d\tilde{\rho}} = \begin{cases} \text{constant}, & \tilde{\rho} \leq \rho_A \\ 0, & \tilde{\rho} > \rho_A \end{cases} \\ \tilde{\rho} &= kn_A r, \quad \tilde{\rho}_A = \rho_A = kn_A r_A, \quad n_A = \text{constant} \end{aligned} \right\} \quad (5.8-4)$$

Here $\tilde{\rho} = \rho_A$ marks the upper boundary of the Airy layer, and n_A is the value of the index of refraction there. From Sections 4.10 and 4.11 we know that the wave equations for this case are given by

$$\frac{d^2(\tilde{\rho}U_l)}{d\tilde{\rho}^2} + \left(\left(\frac{n}{n_A} \right)^2 - \frac{l(l+1)}{\tilde{\rho}^2} \right) (\tilde{\rho}U_l) = 0, \quad \tilde{\rho}V_l = -i \frac{d(\tilde{\rho}U_l)}{d\tilde{\rho}} \quad (5.8-5)$$

$U(\tilde{\rho})$ describes the field along the normal to the plane of propagation (the y -direction in Figure 4-10). For a given value of l , $U(\tilde{\rho})$ also provides the radial spectral component of the electric field, and $V(\tilde{\rho})$ describes the tangential component parallel to the plane of stratification and in the plane of propagation (the θ -direction in Figure 4-10).

This model has been discussed in Section 4.12 where a correspondence between spherical and Cartesian stratification was established, and also in Section 5.7 to obtain an asymptotic form for $G[\rho, \nu]$ valid for $\nu > \rho$. There the solutions to the wave equations in Eq. (5.8-5) are given to a good approximation by the Airy functions with their argument \tilde{y} given by Eq. (4.12-5). Thus, in the medium described by Eq. (5.8-4) for $r \leq r_A$, we have

$$\left. \begin{aligned} \tilde{\rho}U_l &\doteq \{ \text{Ai}[\tilde{y}], \text{Bi}[\tilde{y}] \}, \quad \tilde{\rho}V_l \doteq i|1 - \beta_A|^{1/3} K_{\tilde{\rho}_A}^{-1} \{ \text{Ai}'[\tilde{y}], \text{Bi}'[\tilde{y}] \} n_A \\ \tilde{\rho} &= kn_A r, \quad \beta_A = - \left(\frac{1}{n} \frac{dn}{d\tilde{\rho}} \tilde{\rho} \right)_A \\ \tilde{y} &= - \frac{1}{|1 - \beta_A|^{2/3} K_{\tilde{\rho}_A}} \left((1 - \beta_A)(\tilde{\rho} - \tilde{\rho}_A) - (l - \tilde{\rho}_A) \right) \end{aligned} \right\} \quad (5.8-6)$$

For the region $r > r_A$ the solutions are the spherical Bessel functions of the first and second kinds $\psi_l(\tilde{\rho}), \chi_l(\tilde{\rho})$. We note here that n_A is a constant in the Airy layer. Hence, $-\beta = (dn/d\tilde{\rho})\tilde{\rho}/n = (dn/dr)r/n$.

At the boundary $r = r_A$ we must match these two solution sets for each integer value of l to ensure the continuity of the tangential components of the electromagnetic field across the boundary. We first match the solutions when the $\text{Ai}[\tilde{y}]$ solution applies in the region $r \leq r_A$, that is, the solution that vanishes for decreasing $\tilde{\rho}$ with $\tilde{\rho} < l$. Upon noting from Eq. (5.8-6) that $\partial\tilde{y}/\partial\tilde{\rho} = -|1 - \beta_A|^{1/3} / K_{\tilde{\rho}_A}$ (with $\beta_A < 1$), we set

$$\left. \begin{aligned} \text{Ai}[\tilde{y}_A] &= c_l \psi_l(\tilde{\rho}_A) + d_l \chi_l(\tilde{\rho}_A), \\ -K_{\tilde{\rho}_A}^{-1} (1 - \beta_A)^{1/3} \text{Ai}'[\tilde{y}_A] &= c_l \psi_l'(\tilde{\rho}_A) + d_l \chi_l'(\tilde{\rho}_A) \end{aligned} \right\} \quad (5.8-7)$$

Using the asymptotic forms for both the spherical Bessel functions and the Airy function for $-\tilde{y} \gg 1$ and for $-\hat{y} \gg 1$, we obtain values for the matching coefficients that are given by

$$\left. \begin{aligned} c_l &= \frac{|1 - \beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}} \cos(\tilde{X}_l - \hat{X}_l), \quad d_l = -\frac{|1 - \beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}} \sin(\tilde{X}_l - \hat{X}_l), \\ \tilde{X}_l &= \frac{2}{3}(-\tilde{y}_A)^{3/2} + \frac{\pi}{4}, \quad \hat{X}_l = \frac{2}{3}(-\hat{y}_A)^{3/2} + \frac{\pi}{4} \end{aligned} \right\} \quad (5.8-8)$$

Here \tilde{y} is given in Eq. (5.8-6) and \hat{y} is given by Eq. (5.4-3). Their ratio \hat{y}/\tilde{y} is equal to $|1 - \beta_A|^{2/3}$.

Similarly, for a $\text{Bi}[\tilde{y}]$ solution in the region $r \leq r_A$, that is, the solution that blows up for decreasing $\tilde{\rho}$ with $\tilde{\rho} < l$, we set

$$\left. \begin{aligned} \text{Bi}[\tilde{y}_A] &= e_l \psi_l(\tilde{\rho}_A) + f_l \chi_l(\tilde{\rho}_A), \\ -\frac{|1 - \beta_A|^{1/3}}{K_{\tilde{\rho}_A}} \text{Bi}'[\tilde{y}_A] &= e_l \psi'_l(\tilde{\rho}_A) + f_l \chi'_l(\tilde{\rho}_A) \end{aligned} \right\} \quad (5.8-9)$$

When the asymptotic forms apply, that is, at a location well away from a turning point, we obtain

$$\left. \begin{aligned} e_l &= -\frac{|1 - \beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}} \sin(\tilde{X}_l - \hat{X}_l) = d_l, \\ f_l &= -\frac{|1 - \beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}} \cos(\tilde{X}_l - \hat{X}_l) = -c_l \end{aligned} \right\} \quad (5.8-10)$$

Referring to our discussion in Section 4.6 on incoming and outgoing waves in a Cartesian Airy layer, it follows in this case that outgoing and incoming waves at the boundary can be expressed in the form

$$\text{Ai}[\tilde{y}_A] \mp i \text{Bi}[\tilde{y}_A] = (c_l \mp i d_l) \xi_l^\pm(\rho_A) \quad (5.8-11)$$

The top sign applies to an outgoing wave, and the bottom sign to an incoming wave. Also, the complex coefficients $(c_l \mp i d_l)$ provide the phase delay (modulo 2π) between the incoming and outgoing waves at the boundary due to the refracting medium below. This phase delay offset remains invariant for $r > r_A$ because the medium is taken to be homogeneous for $r \geq r_A$. From Eqs. (5.8-8) and (5.8-10) it follows that the ratio

$$\frac{c_l - id_l}{c_l + id_l} = \exp(2i(\tilde{X}_l - \hat{X}_l)) \quad (5.8-12)$$

provides the round-trip phase delay between the incoming and outgoing wavelets of spectral number l at the boundary $r = r_A$. From Eqs. (5.8-6) and (5.8-8) and noting again that in the Airy layer $\hat{y} / \tilde{y} = |1 - \beta_A|^{2/3}$, it follows that

$$\left. \begin{aligned} \tilde{X}_l - \hat{X}_l &= \frac{2}{3} \left((-\tilde{y}_A)^{3/2} - (-\hat{y}_A)^{3/2} \right) = \\ \frac{2}{3} \cdot \frac{\beta_A}{1 - \beta_A} (-\hat{y}_A)^{3/2} &\doteq \frac{1}{3} \cdot \frac{\rho_A \beta_A}{1 - \beta_A} \left(\frac{2(\rho_A - l)}{\rho_A} \right)^{3/2} \end{aligned} \right\} \quad (5.8-13)$$

We return to this expression momentarily after discussing the form of $G[\rho, \nu]$ for this model of the refracting medium.

Now we use the osculating parameter technique in the spectral representation given in Eq. (5.8-1) to describe the wave in the refracting Airy medium where $r \leq r_A$. The basis functions are the spherical Hankel functions $\{\xi_l^+(\rho), \xi_l^-(\rho)\}$, where $\rho = knr$ and n now is variable and given by Eq. (5.8-4). For a given spectral number the form for the radial term from Eq. (5.8-1) is given by

$$a_l^- \xi_l^- + a_l^+ \xi_l^+ = C_l \left(\exp(-iG[\rho, \nu]) \xi_l^-(\rho) + \exp(iG[\rho, \nu]) \xi_l^+(\rho) \right) \quad (5.8-14)$$

Here C_l is a spectral number-dependent complex factor that depends on boundary conditions, which we will discuss later in regard to turning points; it is not of interest here.

From Eq. (5.7-2) it follows that $G[\rho, \nu]$ for an Airy layer is given by

$$\left. \begin{aligned} G[\rho, \nu] &\doteq \frac{\pi}{6} \left(\frac{1}{n} \frac{dn}{d\rho} \rho \right)_A \left(\Gamma(\hat{y}) - \Gamma(\hat{y}_A) \right), \quad \rho^\dagger(\nu) < \rho \leq \rho_A, \\ G[\rho, \nu] &\rightarrow \frac{2}{3} \left(\frac{1}{n} \frac{dn}{d\rho} \rho \right)_A \left((-\hat{y}_A)^{3/2} - (-\hat{y})^{3/2} \right), \quad \tilde{y}_A, \hat{y} \ll -1 \\ G[\rho, \nu] &\equiv 0, \quad \rho \geq \rho_A \end{aligned} \right\} \quad (5.8-15)$$

where $\Gamma(\hat{y})/3$ is the integral of $g(\hat{y})$ given in Eq. (5.7-1). $\Gamma(\hat{y})$ has been discussed in Section 4.9 with regard to osculating parameters in a Cartesian stratified medium and it is given by Eq. (4.9-5). It is shown in Figure 5-12. $\Gamma(\hat{y})$ has only the two roots shown in this figure and it is monotonic elsewhere.

For a given value of l in both Eqs. (5.8-13) and (5.8-15), what is the value of \hat{y} in Eq. (5.8-15), that is, what value for ρ must we use to match the round-trip phase delay $-2G[\rho, v]$ with the asymptotically exact value of this phase delay $2(\tilde{X}_l - \hat{X}_l)$, given from Eq. (5.8-13)? Noting that n is variable in Eq. (5.8-15), it follows from Eq. (5.8-4) that

$$\left(\frac{1}{n} \frac{dn}{d\rho} \rho \right)_A = - \frac{\beta_A}{1 - \beta_A} \quad (5.8-16)$$

We see upon comparing Eq. (5.8-15) with Eq. (5.8-13) that \hat{y} must be zero. However, the negative argument asymptotic forms for the Airy functions do not apply at $\hat{y} = 0$; instead, we must set $\Gamma(\hat{y}) = 0$ in Eq. (5.8-15). This is discussed in Appendix I, Eq. (I-11) with regard to relating $dG[\rho^\dagger(v), v]/dv$ to $\tilde{\alpha}(v, v)$. The roots of $\Gamma(\hat{y})$ occur at $\hat{y} \doteq -0.2$ and $+0.9$, and $\Gamma(\hat{y})$ reaches a stationary value of almost zero at $\hat{y} = \hat{y}^\dagger$. It is nearly zero there compared to the value of $(-\hat{y}_A)^{3/2}$ when $-\hat{y}_A \gg 1$ (see Figure 5-12). Thus, we set $\rho = \rho^\dagger = v - K_\rho y^\dagger$ to force a near-alignment of Eq. (5.8-15) with Eq. (5.8-13). The relative error of this choice is $\Gamma(\hat{y}^\dagger)/\Gamma(\hat{y}_A) \approx 0.001$ for this example.

In Section 5.5, Eqs. (5.5-22) through (5.5-25), we discussed the spectral coefficients a_l^+ for an outgoing wave in terms of $G[\rho, v]$. The equation for a_l^+ involved a constant of integration $a_l^+[\rho_*, \rho_*]$. The question of the appropriate value of ρ_* to use in this constant of integration was left rather moot in that discussion. Here we set $\rho_* = \rho^\dagger$ and

$$a_l^+[\rho^\dagger, \rho^\dagger] = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG(\rho^\dagger, v)) \quad (5.8-17)$$

This provides a rationale for setting $\rho = \rho^\dagger(v)$ in $2G[\rho, v]$ for the outgoing wave, at least when well away from a turning point so that $\hat{y}_A < 0$. We will return later to this question of linking the incoming and outgoing spectral coefficients.

5.8.3 Numerical Comparisons

Although the close comparison between $2G[\rho^\dagger, v]$ and $(c_l - id_l)/(c_l + id_l)$ at the boundary of the Airy layer is encouraging, we should compare the osculating parameter representation for the wave given in Eq. (5.8-14) with the exact solution obtained from integrating the wave equations in Eq. (5.8-5). Their level of agreement as a turning point is approached is particularly of

interest. Figure 5-13 (a) shows the solution for $U_l(\rho)$, which is the solution to the wave equations in Eq. (5.8-5), and it also shows the osculating parameter solution from Eq. (5.8-14). They are virtually doppelgangers over the range shown for \hat{y} . In this example $\nu = l + 1/2 = \rho_o = 6400k$, $\rho_A = 6432k$, $k = 100 \text{ km}^{-1}$, $n' = 1.56 \times 10^{-4}$ and $n_o = 1$. The solution is closely approximated by $\text{Ai}[\tilde{y}]$. It yields a single-sided bending angle at $\rho_* = \rho_o = 6400k$ of 10 mrad, and this medium has a moderate ray path curvature index β of about 0.1. The general procedure in Figures 5-13 through 5-16 for comparing these solutions is as follows. For initial conditions, we set $U_l(\rho^\dagger) = \text{Ai}[\tilde{y}^\dagger]$ and we also equate their derivatives at this point. This generates a numerical solution for $U_l(\rho)$ that exponentially damps to zero for decreasing ρ in the vicinity of the turning point at $\rho_o = \nu$ (but it blows up for $\rho \ll \rho_o$ because of limited numerical precision). Then at $\rho = \rho_A$, which in the example shown in Figure 5-13 corresponds to a \hat{y} value of about -50 , we set $C_l(a_l^-(\rho)\xi_l^-(\rho) + a_l^+(\rho)\xi_l^+(\rho))$ equal to the numerical solution for $U_l(\rho)$ there and also we equate their derivatives there. This sets the complex value of C_l in Eq. (5.8-14) and gives the osculating parameter solution for decreasing ρ .

Figure 5-13 (b) shows the difference between $U_l(\tilde{\rho})$ and the osculating parameter form with the same initial conditions at the boundary $r = r_A$. Here $G[\rho, \nu]$ is obtained from Eq. (5.5-20), integrated on ρ over the Airy layer from ρ_A to ρ . In the osculating parameter solution for these two panels $G[\rho, \rho_*]$ accumulates about 3 1/2 cycles between r_A and r^\dagger , where $r^\dagger = k^{-1}\rho^\dagger/n(\rho^\dagger)$. This accumulation of 3 1/2 cycles is required to keep the solution $C_l(a_l^-\xi_l^- + a_l^+\xi_l^+)$ in Eq. (5.8-14) aligned in phase with the exact solution $U_l(\rho)$ over the entire Airy layer. Better than 1% numerical agreement holds between solutions except very near a turning point. As expected, for $\hat{y} > \sim -2$ the osculating parameter solution begins to deteriorate, but even at $\hat{y} = 0$ it still is moderately accurate for this example; the difference is 0.007. The differential equations in Eq. (5.8-5) become numerically unstable for $\hat{y} > 0$. Any small numerical errors in matching the boundary conditions or in the numerical integration will magnify greatly in the region $\hat{y} > 0$, that is, $\text{Bi}[\tilde{y}]$ begins leaking into the numerical solution. See the discussion of the connection formula for the WKB solutions in Eq. (5.7-26). Matching boundary conditions closer to $\hat{y} = 0$, for example, at $\hat{y} = -5$ improves the overall agreement, but the osculating parameter solution still deteriorates rapidly for $\hat{y} > 0$. But, the numerical solution for U_l also deteriorates for $\hat{y} > 0$ because of limited precision.

The wavenumber $k = 100 \text{ km}^{-1}$ used in Figure 3-13 is more than 300 times smaller than the wavenumbers of the GPS navigation signals. The smaller value is used here to save computational time and to preserve computational accuracy—for given values of r_A and r_* , the number of cycles in the solution depends linearly on k . It is difficult to maintain coherency between solutions by matching both the solution and its derivative at the same point for highly oscillatory systems. The run-off error is roughly proportional to k . Small errors in the computation of U_l and $G[\rho, v]$ lead inevitably to run-off. An alternate matching strategy to reduce run-off is to pick two separate points and match solutions there, but not their derivatives.

Figure 5-14 shows another comparison between the numerical and osculating parameter solutions. Here a moderately strong refractivity gradient has been used, leading to single-sided bending angle variations of more than 30 mrad over a tangency point altitude range of 8 km. Panel (a) shows the profile of the refractivity change, and panel (b) shows the resulting bending angle profile. In the lowest layer, $\rho dn/d\rho = -0.8$, which gives a ray path curvature index of $\beta = 4/9$. But in the middle layer $\rho dn/d\rho = +0.8$, which yields $\beta = -4$, a rather extreme negative ray path curvature. Panel (c) compares the two solutions, $U_l(\rho)$ and $a_l^-(\rho)\xi_l^-(\rho) + a_l^+(\rho)\xi_l^+(\rho)$. Here the radial distance r is used for the abscissa instead of \hat{y} with a range of 8 km, and with $k = 400 \text{ km}^{-1}$. Per the discussion in Section 5.7 concerning Eqs. (5.7-25) through (5.7-26), by forcing the osculating parameter solution to follow the Airy layer solution for \hat{y} values greater than ~ -2 , roughly 1/2 km above the turning point for $k = 400 \text{ km}^{-1}$ ($\sim 30 \text{ m}$ for GPS wavelengths), one can greatly improve the solution below the turning point, until the numerical solution itself begins to fail.

Figure 5-15 compares the wave equation solutions for a severely refracting medium that includes a super-refracting layer, 1 km thick in the range $r_d < r < r_u$. The refractivity profile is shown in panel (a). Within the super-refracting layer the ray path curvature index has a value of $\beta = 2$. Above the layer, $\beta = 0.4$, and below, $\beta = 0.3$. These refractive gradients lead to enormous swings in the ray path bending angle. Panel (b) shows the resulting single-sided bending angle profile versus tangency point radius, including the super-refracting zone $r^c < r_* < r_u$. For tangency points within this range no rays can occur when spherical symmetry applies. Panel (c) shows the variation in impact parameter in the vicinity of this super-refracting layer. The impact parameter has a negative slope within the layer. Panel (d) compares the wave equation solutions. In this example the turning point is well below the critical radius, i.e., $r_o < r^c$. Good agreement holds except near the turning point. A wavenumber value of $k = 1000 \text{ km}^{-1}$ is used in Figures 5-15 and 5-16. The

wave equation solutions in Figure 5-15 (d) show a slight amplitude bulge and a slight reduction in frequency in the vicinity of the super-refracting layer, which is expected. Referring to the WKB solution connection formulas in Eq. (5.7-26), the amplitude and frequency factor expressed in length units is $f_l(\rho) \doteq -k^2 \hat{y} / K_v^2 \approx -46 \hat{y} \text{ km}^{-2}$. When $\beta < 1$, $\partial \hat{y} / \partial r < 0$; but in a super-refracting medium where $\beta > 1$, then $\partial \hat{y} / \partial r > 0$. It follows that $\partial f_l / \partial r > 0$ for $r < r_d$; but within the super-refracting layer, $r_d < r < r_u$, $\partial f_l / \partial r < 0$. For a turning point below the critical tangency point, i.e., for $v < kn(r^c)r^c$, $f_l(\rho)$ remains positive for all $r > k^{-1}v / n(r)$, but it becomes less positive with increasing r within the super-refracting layer, which is reflected in the figure by the increasing amplitude ($\sim 1/\sqrt[4]{f_l}$) and decreasing frequency ($\sim \sqrt{f_l}$) within the layer. Above the layer $\partial f_l / \partial r$ returns to positive territory. When $v < kn(r^c)r^c$, the l th spectral component of the wave powers through the super-refracting layer, as shown in Figure 5-15 (d), just as the corresponding ray, with an impact parameter value $\rho_* = v \leq kn(r^c)r^c = kn(r_u)r_u$, also does, albeit severely refracted as $\rho_* \rightarrow kn(r^c)r^c$ from below.

As the ray path tangency point approaches the critical radius $r_* = r^c$, the bending angle approaches a limit, shown in panel (b), and the wave solution shows a marked transition from an oscillatory form to an exponential form in the vicinity of the layer. A delicate situation occurs when the turning point is such that $kn(r^c)r^c \leq v \leq kn(r_d)r_d$. Here $f_l(\rho)$ reverses sign at $kn(r)r = v$, with the lower root to the equality lying in the range $r^c < r < r_u$, and the upper root at $r > r_u$. In this case we have two zones where $U_l(\rho)$ must be exponential-like: in that region in and above the super-refracting layer where $f_l < 0$, and below the turning point where $f_l < 0$ again. Inasmuch as there are no rays for tangency points in the range $r^c < r_* < r_u$, we expect the wave equation equations to give a damped amplitude for the field.

Lastly, Figure 5-16 shows the case where the spectral number has been increased (actually, the refractivity profile was lowered in altitude relative to a fixed turning point) so that the spectral number lies in the range $kn(r^c)r^c \leq v = kn(r_o)r_o \leq kn(r_d)r_d$. These panels show the wave equation solutions and the profile for \hat{y} . Panels (a) and (b) are for the case where $v = kn(r^c)r^c = kn(r_u)r_u$, the critical impact parameter value marking the boundaries of the zone corresponding altitude range $r^c < r_* < r_u$ within which no ray path tangency points may lie. Panels (c) and (d) are for the case where the spectral number lies within the critical range $kn(r_u)r_u < v < kn(r_d)r_d$, the

equivalent of about 150 m in radial distance above r^c for $k = 1000 \text{ km}^{-1}$. Panels (a) and (c) were normalized so that the amplitudes of all of the wave equation solutions in (a) and (c) have the same asymptotic value with increasing r well above the atmosphere. Here the agreement between $U_l(\rho)$ and the osculating parameter solutions degrades significantly in the vicinity and below the super-refracting layer because \hat{y} becomes positive over a limited interval about the upper boundary at $r = r_u$. It also becomes positive again below the turning point. Although the amplitudes of the two solutions diverge below the layer, their phases remain aligned right down to the turning point.

A super-refracting layer acts like a second turning point when the spectral number lies in the critical zone $kn(r^c)r^c \leq \nu \leq kn(r_d)r_d$. We note from Eq. (5.7-26) that the slightest hint of a positive value for \hat{y} in the super-refractive zone ($\sqrt{-f_l} \doteq k\sqrt{\hat{y}}/K_v \approx 7\sqrt{\hat{y}}$) causes $U_l(\rho)$ to damp to near-zero below the lower critical point where $\hat{y} = 0$. The easiest way to see this is to consider the reverse sense of propagation. Consider a solution for $U_l(\rho)$ that is forced to damp to zero below the turning point at $\rho = \rho_o$. With these initial conditions on $U_l(\rho)$ and its derivative, the solution is then propagated outward. When the region above the turning point where $f_l(\rho)$ first becomes negative is encountered, the solution becomes exponential-like. This causes the amplitude of $U_l(\rho)$, which still has an Airy function-like character (of the first kind) below this point, to be catapulted to an astronomical magnitude if $f_l(\rho)$ remains negative for a sufficient interval³. But we have normalized the asymptotic values of the outbound amplitudes to correspond to the amplitude of an incoming wave, which is essentially invariant over the narrow range of spectral numbers considered here. This forces the amplitude of $U_l(\rho)$ in the region below the lower point at which $f_l(\rho) = 0$ to be greatly diminished when $\hat{y}_{\text{MAX}} = \hat{y}_u$ is positive. In panel (a), $\hat{y}_{\text{MAX}} \doteq 0$, and the amplitude of $U_l(\rho)$ below the layer is beginning to attenuate. But in panel (c), \hat{y}_{MAX} is barely positive, 1.1, and $f_l(\rho)$ is negative only over a 40 m interval

³ An easy refractivity model with which to see this amplification process is given by $n' = \text{constant} < -157 \times 10^{-6} \text{ km}^{-1}$ in the super-refracting layer $r_d \leq r \leq r_u$, and $n' \equiv 0$ elsewhere. The wave equation solutions are, below the layer: $\text{Ai}[\hat{y}]$, within the layer: $a\text{Ai}[\hat{y}] + b\text{Bi}[\hat{y}]$, and above the layer: $c\text{Ai}[\hat{y}] + d\text{Bi}[\hat{y}]$. Equating these solutions and their derivatives across their respective boundaries, at $r = r_d$, and at $r = r_u$ where $\hat{y} = \hat{y}_{\text{max}}$, leads to an explicit evaluation of the coefficients c and d . It can be shown that the amplitude $|c^2 + d^2|^{1/2} \rightarrow 2(\beta - 1)^{1/6} \exp(2(\beta/(\beta - 1))\hat{y}_{\text{max}}^{3/2}/3)$, when $\hat{y}_{\text{max}} > 0$. The extreme case $n' \rightarrow -\infty$ with Δn finite is addressed in Chapter 3 (Mie scattering theory) and in Section 5.13.

about $r = r_u$. Already for these slight intrusions the amplitude of $U_l(\rho)$ is an order of magnitude diminished below the layer. For GPS wavelengths, $\Delta\hat{y} = 1$ corresponds to $\Delta r = 14$ m and $\sqrt{-f_l} \approx 70\sqrt{\hat{y}}$. Therefore, the cutoff in amplitude of $U_l(\rho)$ below a super-refracting layer for spectral numbers greater than the limiting value yielding $\hat{y}_{\text{MAX}} = 0$ is extremely sharp.

The lower point for $r < r_u$ in the super-refracting layer where $\hat{y} = 0$ is also the first point where the bending angle integral for a hypothetical ray with an impact parameter value of $\rho_* = v$ would become complex, if such a ray were to exist. Here $kn(r)r$ becomes smaller than ρ_* for increasing r above this point until the upper point where $\hat{y} = 0$ is reached at $r < r_u$.

A better approximation approach (but not used here) for the osculating parameter solution to obtain closer alignment with $U_l(\rho)$ in and below the super-refraction zone is to break the spherical medium into regimes. These are designated by the Roman numerals in panel (d), which have boundaries in the radial coordinate r at the points where $\hat{y} \approx -2$. In I and IV the osculating parameter solution applies because $\hat{y} < -2$ throughout these regimes. In the tunneling regimes, II, III and V, the refractivity profile is approximated by an Airy layer, each of which has a wave equation solution given by Eq. (5.8-6). Since \hat{y} need not be too positive before the solution below the layer is essentially damped to zero, this Airy layer approximation should be valid. The coefficients of these different solutions are tied together by using the continuity conditions on the solutions and their derivatives across each boundary; these boundaries are marked in (d). The amplitude of the osculating parameter solution in IV will then damp to zero rapidly below the layer, which $U_l(\rho)$ also does, except for spectral numbers in the transition zone where $\hat{y}_{\text{MAX}} \approx 0$. On the other hand, in region IV below the super-refraction zone, $G[\rho, v]$ contains valid phase information because $\hat{y} < -2$.

In summary, Figures 5-13 through 5-16 show good agreement between the exact and osculating parameter solutions over almost all regions except those in the immediate neighborhood of a turning point or in that delicate transition across the critical spectral number range $v \approx kn(r^c)r^c$, below a super-refracting layer. Here $\hat{y}_{\text{MAX}} \approx 0$, but not yet positive enough to rapidly damp $U_l(\rho)$ to zero below the layer. In calculating the electric field from the integrals over spectral number that are given in Eq. (5.8-1b), Section 3.16 shows that the principal contribution to these integrals comes from a neighborhood in spectral number where the phasor in the integrand is varying the least. This is a stationary phase neighborhood, which may or may not be unique, depending on the refractivity profile. If the field is being evaluated at an incoming point well away from a turning point, then a stationary phase neighborhood will not be located near $\hat{y} = 0$, but rather in negative territory. Therefore, except for

incoming positions near a turning point, we expect the accuracy of the osculating parameter technique applied to the spectral integrals in Eq. (5.8-1b) to be adequate. Moreover, near a turning point one can use the Airy layer approximation to greatly improve the accuracy there, which is discussed in Section 5.11. The issue of outgoing points remains, and it is discussed later.

5.8.4 Comparison of Phase Delays in an Airy Layer from Wave Theory and Geometric Optics

According to geometric optics, the single-sided phase delay for an initially collimated incident ray with an impact parameter $\rho_* \leq \rho_A$ is given by

$$ck\tau = k \int_{r_*}^{r_A} \frac{n^2 r dr}{\sqrt{n^2 r^2 - n_*^2 r_*^2}} = - \int_{\rho_*}^{\rho_A} \frac{d \log n}{d\rho} \sqrt{\rho^2 - \rho_*^2} d\rho + \left\{ \begin{array}{l} \sqrt{\rho_A^2 - \rho_*^2} + \rho_* (\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_A, \rho_*)) \end{array} \right\} \quad (5.8-18)$$

The second equality follows from integrating by parts and using Eq. (5.6-2) for the bending angle. Using Bouguer's law, we see that the bottom line simply equals $\rho_A \cos \theta_A \sec(\Delta \tilde{\alpha})$ through second order in $\Delta \tilde{\alpha} = \tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_A, \rho_*)$. Therefore, the extra single-sided phase delay caused by the refractive gradient for a ray with an impact parameter $\rho_* \leq \rho_A$ is given by

$$ck\Delta\tau = - \int_{\rho_*}^{\rho_A} \frac{d \log n}{d\rho} \sqrt{\rho^2 - \rho_*^2} d\rho \doteq \frac{\beta_A \rho_*}{3(1 - \beta_A)} \left(\frac{2(r_A - r_*)}{r_*} \right)^{\frac{3}{2}} \quad (5.8-19)$$

The approximate expression on the RHS applies to the Airy layer model and it is accurate when thin atmosphere conditions apply and $r_*/\lambda \gg 1$ (See Appendix A). This expression also gives about 3 1/2 cycles for the conditions shown in Figure 5-13 and it matches the phase delay expression given in Eq. (5.8-13).

Eq. (5.8-13), which is an asymptotically exact result for an Airy layer, Eq. (5.8-15), which gives $G[\rho, v]$ from use of the osculating parameter technique, and Eq. (5.8-19), which is from geometric optics, all essentially agree on the phase induced by the refractive gradient on a wave passing through the Airy layer. We return to this model in Section 5.11 where calculating the field at a turning point is discussed.

Eq. (5.8-19) provides us with further insight into the character of $G[\rho, v]$. When $v^* = \rho_*$, $G[\rho, \rho_*]$ provides the extra path delay for a ray with an impact parameter value of ρ_* , both from the curvature component of the ray that is

induced by the refractive gradient and from the change in wave velocity along the ray in the medium.

5.8.5 Asymptotic Matching the Spectral Coefficients for Incoming and Outgoing Waves

Let (r_*, θ_*) mark a turning point and $\rho_* = kr_* n_*$ is the impact parameter associated with the ray passing the turning point. We know from Bouguer's law for a spherical symmetric medium that ρ_* is a constant when evaluated along that ray path.

When the planar asymptotic boundary conditions apply to the approaching wave, then the incoming spectral coefficient $a_l^-(\rho)$ is uniquely determined from Eq. (5.5-21). It is independent of the impact parameter ρ_* associated with the point (r, θ) at which $a_l^-(\rho)$ is evaluated because the asymptotic boundary conditions in Eq. (5.5-3) for a planar approaching wave are independent of impact parameter. Because the wave front surface is not symmetric about the turning point boundary (the approaching waves are collimated prior to impacting the medium; the departing waves are dispersed), the outgoing coefficient depends on the value of ρ_* . For the outgoing wave it follows from Eq. (5.5-23) that we can obtain $a_l^+[\rho, \rho_*]$ if we know the value of $a_l^+[\rho, \rho_*]$.

Consider first the symmetric problem mentioned earlier where the electromagnetic wave is planar along the line $\theta = \pi/2$, $\phi = 0$, i.e., along the x axis in Figure 4-10. From Bauer's identity in spherical coordinates we have along the x -axis

$$1 = \sum_{l=1}^{\infty} i^l (2l+1) \left(\frac{\psi_l(\rho)}{\rho} \right) P_l(0) \quad (5.8-20)$$

which holds for all values of ρ . As the wave propagates away from the line $\theta = \pi/2$, which is along the x axis, the cophasal normal path in the plane $\phi = 0$ will depend on its initial position ρ along the x axis from which it started. Thus, for every value $\rho = \rho_*$, the spectral coefficients must have the form

$$a_l^{\pm}[\rho, \rho_*] = i^{l-1} \frac{2l+1}{l(l+1)} \exp(\pm i(G[\rho, v] - G[\rho_*, v])) \quad (5.8-21)$$

Using these forms for the spectral coefficients in the scalar potentials given in Eq. (5.8-1), it is readily shown with the Helmholtz equation in Eq. (5.2-6) that the radial component of the electric field is invariant along the x axis, and given by the RHS of Eq. (5.8-20).

To convert this symmetric form for the electromagnetic wave into a non-symmetric form, that is, the version where the wave is asymptotically

collimated on the incoming side ($\pi > \theta > \theta_*$), we merely multiply both a_l^- and a_l^+ in Eq. (5.8-21) by $\exp(-iG[\rho_*, v])$ (thus preserving their equality at $\rho = \rho_*$). We will show later that this is equivalent to rotating clockwise each cophasal normal path by the angle $\tilde{\alpha}(\rho_*, \rho_*)$, where ρ_* is the impact parameter of the path. For this case as $\rho \rightarrow \infty$

$$\left\{ \begin{array}{l} a_l^- \rightarrow i^{l-1} \frac{2l+1}{l(l+1)} \\ a_l^+ \rightarrow i^{l-1} \frac{2l+1}{l(l+1)} \exp(-2iG[\rho_*, v]) \end{array} \right\}, \quad \rho \rightarrow \infty \quad (5.8-22)$$

The question arises as to what value to use for ρ_* in these expressions in either Eq. (5.8-21) or Eq. (5.8-22). Consider the variability of $a_l^-(\rho)$ with ρ for the collimated case when the spectral number is set to a fixed value v_o . At a point (r, θ) on the approaching side at large distances where $n \rightarrow 1$ we set $v_o = kr \sin \theta$; that is, v_o becomes the impact parameter for the cophasal normal passing through the point (r, θ) . From Eq. (5.5-21) the evolution of $a_l^-(\rho)$ with v while travelling along this particular ray path can be obtained by studying the behavior of $G[\rho, v]$ with v fixed. Figure 5-6 shows an example of $G[\rho, v_o]$ versus ρ in the close vicinity of $\rho = v_o$. $G[\rho, v]$ has a stationary value at $\rho = v_o - \hat{y}^+ K_\rho$, where \hat{y}^+ is given by Eq. (5.7-10). This point marks a stationary phase point for $a_l^-(\rho)$ (and for $a_l^+(\rho)$) with respect to ρ . We know that the forms for $G[\rho, v]$ given in Eq. (5.5-20) or in Eq. (5.7-2) begin to fail for increasing $\hat{y} > 0$, and that the correct form for $G[\rho, v]$ rapidly assumes a constant value. Figure 5-7 shows a comparison between the exact phase rate $\partial \vartheta_l^- / \partial \rho$ and $\partial G[\rho, v] / \partial \rho$ for an Airy layer.

Consistent with the approximation chosen in Eq. (5.7-27) for $G[\rho, v]$, we set $a_l^+ = a_l^-$ at this point $\hat{y} = \hat{y}^+$ to ensure no singularity at the origin from the Hankel functions and to attain a close match with the refractive bending angle. Note that this stationary phase point in ρ space varies with the value of v_o . At this stationary point, $\rho^\dagger = \rho^\dagger(v) = v - K_{\rho^\dagger} \hat{y}^+$, we set

$$a_l^+[\rho^\dagger, \rho^\dagger] = a_l^-[\rho^\dagger] = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho^\dagger, v]) \quad (5.8-23)$$

This is exactly the same form for $a_l^+[\rho_*, \rho_*]$ that we obtained in Eq. (5.8-17) at the boundary of an Airy layer to force alignment between the roundtrip delay

obtained from the osculating parameter technique with the roundtrip delay from the Airy layer solution. Here $G[\rho, v]$ would include the delay from the Airy layer and from the overlying medium. Of course, our Airy layer can be made as thin as we please relative to the extent of the overlying medium. Thus, we asymptotically match the incoming and outgoing spectral coefficients.

From Eq. (5.5-23) it follows that $a_l^+(\rho)$ is given by

$$a_l^+(\rho) = a_l^+[\rho^\dagger, \rho^\dagger] \exp(-i(G[\rho^\dagger, v] - G[\rho, v])) = \left\{ \begin{array}{l} i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i(2G[\rho^\dagger(v), v] - G[\rho, v])) \end{array} \right\} \quad (5.8-24)$$

At the LEO, which is located at (r_L, θ_L) and assumed to be outside of the refracting medium where $n \equiv 1$, the second term in Eq. (5.8-24) is zero, that is, $G[\rho_L, v] \equiv 0$. It follows that

$$a_l^+(\rho_L) = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho^\dagger(v), v]), \quad \rho^\dagger = v - y^\dagger K_{\rho^\dagger} \quad (5.8-25)$$

Thus, $-2G^\dagger(v) \equiv -2G[\rho^\dagger(v), v]$ is the two-way phase accumulation of the l th spectral coefficient from the stationary point $\rho^\dagger = kr^\dagger n(\rho^\dagger)$ to outside of the refracting medium. *It corresponds in geometric optics to the extra phase accumulation induced by the refractivity gradient while travelling along a ray that has completely transected the atmosphere and that has an impact parameter value of v .*

We will show later that the value of v that yields a stationary phase in the wave theory spectral representation in Eq. (5.8-1) is essentially equal to the value of the impact parameter for the ray passing through the LEO provided that the inequality $\rho_* |d^2 n / d\rho_*^2| \ll \tilde{\alpha}(\rho_*, \rho_*)$ is satisfied. The latter value is given from Bouguer's law by

$$v = \rho_* = kr_* n_* = kr_L \sin(\theta_L + \alpha_L(\rho_*)) \quad (5.8-26)$$

where $\alpha_L(\rho_*) = 2\tilde{\alpha}(\rho_*, \rho_*)$ is obtained from Eq. (5.6-5). When multipath situations occur, this value may not be unique; we will discuss that later. It follows that the stationary point for the spectral number in wave theory is given by

$$v = v^\dagger = \rho_* + \hat{y}^\dagger K_{\rho_*} \quad (5.8-27)$$

A continuing issue for outgoing points concerns the accuracy of the adopted form in Eq. (5.7-27) for $G[\rho, v]$, and therefore, the accuracy of

$-2G[\rho^\dagger(v), v]$ in representing the round-trip phase delay. Because of the failure of the osculating parameter technique for decreasing $\rho < \rho^\dagger$, this adopted form has some error. But the adopted form for $G[\rho, v]$ does have a stationary value at $\rho = \rho^\dagger$, and the correct form must rapidly approach a fixed value for $\rho < \rho^\dagger$ (see Figure 5-7). In addition, we have already noted in Section 5.7 the close correspondence between $G[\rho^\dagger(v), v]$ and the component of the phase delay in the eikonal equation induced by the refractivity gradient for a ray with an impact parameter value of v . Also noted there and in Appendix J is the high accuracy of the relationship, $dG[\rho_*(\rho_*), \rho_*] / d\rho_* \doteq \tilde{\alpha}(\rho_*, \rho_*)$, provided that the curvature in the refractivity profile is moderate. In this section, the close comparison between $-2G[\rho^\dagger(v), v]$ and the exact solution for the round-trip phase delay in an Airy layer has been noted. Moreover, a byproduct of the numerical solutions presented in Figures 5-13 through 5-16 (along with the second numerical solution for $U_l(\rho)$ with the boundary conditions generated from $\text{Bi}[\tilde{y}^\dagger]$) is the exact value of the extra phase delay induced by the refractivity gradient of the medium. $G[\rho^\dagger(v), v]$, based on Eq. (5.7-27) and calculated from Eq. (5.7-2), may be directly compared to this numerical result for the different refractivity models assumed in these examples. The agreement is generally in the range 0.5-0.1% when the curvature index β is not too close to unity. The computational imprecision in these results also is of the order of 0.1%. Section 6.5 compares the adopted form for $dG[\rho^\dagger(v), v] / dv$, a key spectral quantity in the recovery of the refractivity profile from the LEO amplitude and phase observations, with the exact form for an Airy layer. It gives the error as a function of the ray path curvature index.

An alternative rationale for picking the form given in Eq. (5.8-24) for a_l^+ is based on probabilistic arguments derived from summing over all possible ray paths. The rationale is similar to the Feynman sum-over-histories technique in quantum electrodynamics to calculate the probability of a quantum event. On the incoming side the probability density distribution of impact parameter values for the rays is a flat curve; each value is equally likely to occur. Bouguer's law requires that the flat distribution be preserved after atmospheric encounter. Therefore, the values of ρ_* at a turning point are uniformly distributed. If we set the outgoing spectral coefficient to be the spatial average over all possible impact parameter values to be used in Eq. (5.8-22), we end up with an averaging integral to evaluate. The stationary phase value of this integral yields $\rho_* = \rho^\dagger(v)$. Therefore, the stationary value of $G[\rho, v]$ at $\rho = \rho^\dagger(v)$ should be adopted in Eq. (5.8-22).

5.9 Interpreting Wave Theory in a Refracting Medium Using the Stationary Phase Technique

We now return to the wave theory spectral representation in Eqs. (5.8-1ab) for the electric field at a point (r, θ) in the refracting medium. Before presenting numerical examples using this spectral representation, we apply stationary phase concepts to aid in the interpretation of those numerical results to compare with geometric optics. We follow Chapter 3 closely here; in particular, we refer to the material in Sections 3.10 through 3.13. The integrals in Eq. (5.8-1b) are characterized by a slowly varying factor multiplied by the sum of four phasors that are rapidly oscillating over most of the spectral number space. The main contributions to these integrals come from neighborhoods where any one of these phasors is varying the least. Our task now is to use the stationary phase technique on these integrals to identify the possible stationary phase neighborhoods for each phasor and to calculate the values for E_r and E_θ at a given point (r, θ) .

5.9.1 Geometric Interpretation of the Phasors

We rewrite the asymptotic forms for the Hankel and Legendre functions in Eq. (5.8-1) in a phasor form that provides a useful geometric interpretation. Except for a point (r, θ) located very near a turning point, we will show that the stationary phase neighborhoods in v -space are sufficiently below the value $\rho = krn(\rho)$ so that the negative argument asymptotic forms for the Airy functions given in Eq. (3.8-7) can be used. In this case, from Eqs. (3.10-1) through (3.10-4), the spherical Hankel functions can be rewritten as

$$\left. \begin{aligned} \xi_l^\pm(\rho) &\sim \left(\frac{\rho}{D_v}\right)^{\frac{1}{2}} \exp\left(\pm i\left(D_v + v(\theta_v - \pi/2) - \pi/4\right)\right), \\ \xi_l^{\pm'}(\rho) &\sim \left(\frac{D_v}{\rho}\right)^{\frac{1}{2}} \exp\left(\pm i\left(D_v + v(\theta_v - \pi/2) - \pi/4\right)\right) \end{aligned} \right\} \quad (5.9-1)$$

where

$$D_v = \sqrt{\rho^2 - v^2}, \quad \theta_v = \sin^{-1}\left(\frac{v}{\rho}\right), \quad \rho = krn(\rho), \quad v < \rho \quad (5.9-2)$$

Similarly, from Eq. (3.10-5) the asymptotic form for the Legendre polynomial $P_l^1(\cos \theta)$ is given by

$$P_l^1(\cos \theta) \sim -\sqrt{\frac{l}{2\pi \sin \theta}} \left(\exp \left[i \left(v\theta + \frac{\pi}{4} \right) \right] + \exp \left[-i \left(v\theta + \frac{\pi}{4} \right) \right] \right) \quad (5.9-3)$$

The geometric interpretation of θ_v as the central angle to the point (ρ, θ_v) and D_v as its tangential distance in phase units from a spherical caustic surface of radius v has been given in Chapter 3, Eqs. (3.11-1) and (3.11-2), and in Figures 3-13 and 3-14. Those figures are applicable to an outgoing wave, but the concepts are the same, whether incoming or outgoing. Figure 3-13 shows the geometric relationships given by Eq. (5.9-2). Two rays, originally collimated from the GPS direction, reach the point L located at (ρ, θ) . The direct path is a straight path to the tangential point P_1 and then straight on to the point L. Along the retrograde path via P_2 , the ray arrives tangentially at P_2 , travels along the arc $P_2 P'_2$ and then departs tangentially from P'_2 going straight on to the point L. We note that when it is assumed that the stationary phase value v^* lies in the range $0 < v^* < \rho$, it follows that $0 < \theta_{v^*} < \pi/2$. As v increases through its range of values in the spectral integrals in Eq. (5.8-1b), the radius of this caustic sphere expands and its center descends.

The four phasors appearing in the integrals in Eq. (5.8-1b) result from the product of the spherical Hankel and Legendre functions in Eqs. (5.9-1) and (5.9-3) times the incoming and outgoing spectral coefficients given in Eqs. (5.5-21) and (5.8-24), respectively. To further interpret these phasors geometrically we let

$$\left. \begin{array}{l} \text{outgoing: } \theta = \theta_v + \Delta\theta_v, \\ \text{incoming: } \theta = \pi - (\theta_v + \Delta\theta_v) \end{array} \right\} \quad (5.9-4)$$

Because the stationary neighborhoods for v will be close to $v^* = \rho \sin(\theta + \alpha)$, we would expect $\Delta\theta_v$ to be small in these neighborhoods. Figure 3-14 provides a geometric interpretation of $\Delta\theta_v$ in terms of the extra phase $v\Delta\theta_v$ along the $\theta = 0$ direction that results from the offset $\theta - \theta_v$ for an outgoing wave (and $\theta - (\pi - \theta_v)$ for an incoming wave).

5.9.2 Stationary Phase Conditions

We now insert the asymptotic forms in Eqs. (5.9-1)-(5.9-3) into Eq. (5.8-1b), and we substitute the forms for the spectral coefficients a_l^- and a_l^+ given by Eqs. (5.5-21) and (5.8-24), respectively. After some manipulation of Eq. (5.8-1b), the spectral representation for $E_r(r, \theta)$ and $E_\theta(r, \theta)$ becomes

$$(5.9-5)$$

$$\left. \begin{aligned} E_r &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty \left(\frac{\sin^3 \theta_v}{\cos \theta_v} \right)^{\frac{1}{2}} (e^{i\Psi(+,+)} + e^{i\Psi(+,-)} + e^{i\Psi(-,+)} + e^{i\Psi(-,-)}) dv \\ E_\theta &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty \left(\frac{\sin 2\theta_v}{2} \right)^{\frac{1}{2}} (e^{i\Psi(+,+)} + e^{i\Psi(+,-)} + e^{i\Psi(-,+)} + e^{i\Psi(-,-)}) dv \end{aligned} \right\}$$

where the four phases are given by

$$\left. \begin{aligned} \Psi(+,+) &= D_v + v(\theta_v + \theta) + \pi/4 + G[\rho, v] - 2G[\rho^\dagger, v] \\ \Psi(+,-) &= D_v + v(\theta_v - \theta) - \pi/4 + G[\rho, v] - 2G[\rho^\dagger, v] \\ \Psi(-,+) &= -D_v - v(\theta_v - \theta - \pi) + 3\pi/4 - G[\rho, v] \\ \Psi(-,-) &= -D_v - v(\theta_v + \theta - \pi) + \pi/4 - G[\rho, v] \end{aligned} \right\} \quad (5.9-6)$$

Here (+, -) designates use of the spherical Hankel function of the first kind ξ_l^+ and the negative sign in the exponential for the Legendre polynomial; similarly for the other combinations. In the integrals in Eq. (5.9-5), θ_v varies slowly with v , but the phasor terms are rapidly varying except at stationary phase points. The principal contribution to these integrals comes from neighborhoods about stationary phase points. In the stationary phase technique the phase of each phasor is expanded in a Taylor series through second degree about possible stationary phase values of v . Thus, the first degree term is zero and the Taylor series contains only a zeroth degree term and a quadratic term. Upon evaluating the slowly varying terms at a stationary phase point, the integral reduces to a Fresnel integral.

To see if a stationary phase point exists for the four phasors given in Eq. (5.9-6) we substitute Eq. (5.9-4) into Eq. (5.9-6) to obtain

$$\left\{ \begin{aligned} \Psi(+,+) &= D_v + v(2\theta_v + \Delta\theta_v) + \pi/4 + G[\rho, v] - 2G[\rho^\dagger, v] \\ \Psi(+,-) &= D_v - v\Delta\theta_v - \pi/4 + G[\rho, v] - 2G[\rho^\dagger, v] \\ \Psi(-,+) &= -D_v - v(2\theta_v - 2\pi - \Delta\theta_v) + 3\pi/4 - G[\rho, v] \\ \Psi(-,-) &= -D_v + v\Delta\theta_v + \pi/4 - G[\rho, v] \end{aligned} \right. \quad (5.9-6')$$

Comparing the terms in Eq. (5.9-6') with Figure 3-13, we conclude (+, +) is associated with an outgoing wave from the far side of the scattering sphere, that is, it is associated with the retrograde path; (-, -) is associated with the incoming wave on the near side, and so on. We can eliminate by inspection the phasors $\Psi(+, +)$ and $\Psi(-, +)$, because these are associated with waves that travel around the far side of the sphere in Figure 3-13 via point P_2 . They will

provide negligible contributions to the spectral integrals for the electric field when r_o / λ is large and the point L is located well into the first or second quadrants in θ . The term $2\nu\theta_v$ in their arguments results in very high rates of phase accumulation for essentially all values of ν .

5.9.3 Stationarity of $\Psi(\pm, \pm)$

Returning to Eq. (5.9-6) we take the partial derivative of the phase $\Psi(\pm, \pm)$ with respect to ν in each of these four phasor combinations that appear in Eq. (5.9-6) and attempt to set the resultant equal to zero. We conclude (remembering that $0 < \theta_v \leq \pi/2$ and that $0 < \theta < \pi$) the following with regard to the possibility of achieving a stationary value for each of these phases:

$$\partial\Psi(+, +) / \partial\nu = 0:$$

$$\text{no, if } \theta > \partial(2G[\rho^*, \nu] - G[\rho, \nu]) / \partial\nu,$$

$$\partial\Psi(+, -) / \partial\nu = 0:$$

$$\text{yes, if } \theta < \pi/2 - \partial(2G[\rho^*, \nu] - G[\rho, \nu]) / \partial\nu,$$

$$\partial\Psi(-, +) / \partial\nu = 0:$$

$$\text{no, if } \theta \geq 0$$

$$\partial\Psi(-, -) / \partial\nu = 0:$$

$$\text{yes, if } \theta > \pi/2 - \partial G[\rho^*, \nu] / \partial\nu.$$

As already mentioned, for an occultation from an LEO we can effectively rule out the $(+, +)$ and $(-, +)$ combinations. These are contributions to the integrals in Eq. (5.9-5) that originate from the far-side of the sphere (Figure 3-13). They are negligible when θ lies well into the upper quadrants and when the ratio r_o / λ is very large, both of which are assumed here. The combination $(+, -)$ corresponds to an outgoing wave on the near-side of the sphere where the spherical Hankel function $\xi_l^+(\rho)$ is used; the combination $(-, -)$ corresponds to an incoming wave on the near-side where $\xi_l^-(\rho)$ applies.

5.9.4 Plane Waves

As an illustrative case, consider the stationary phase possibilities for a planar wave in a homogeneous medium. Here $n' \equiv 0$ and therefore, $G[\rho, \nu] \equiv 0$. An appropriate spectral representation for this case is given from Bauer's identity (Section 5.3) in spherical coordinates

$$\exp(i\rho \cos \theta) = \sum_{j=0}^{\infty} i^l (2l+1) \frac{\psi_l(\rho)}{\rho} P_l(\cos \theta) \quad (5.9-7)$$

Here, expressing $i^l \psi_l P_l$ in terms of phasors (with $\psi_l = (\xi_l^+ + \xi_l^-)/2$) leads to the same combinations $\Psi(\pm, \pm)$ given in Eq. (5.9-6) but with $G[\rho, v] \equiv 0$. Attempting to set $\partial \Psi(\pm, \pm) / \partial v = 0$ for the four possible sign combinations in the region $0 < \theta < \pi$ produces the following conditions on θ_v :

$$\begin{aligned} (+, +): \theta_v &= -\theta; \text{ impossible} \\ (+, -): \theta_v &= \theta; \text{ possible if } 0 < \theta < \pi/2 \\ (-, +): \theta_v &= \theta + \pi; \text{ impossible} \\ (-, -): \theta_v &= \pi - \theta; \text{ possible if } \pi > \theta > \pi/2 \end{aligned}$$

It follows that the $(+, -)$ case corresponds to a departing planar wave and that the $(-, -)$ case corresponds to an approaching wave. In either case a stationary value for $\Psi(\pm, -)$ is achieved when the spectral number $v = v^* = \rho_* = \rho \sin \theta$. For this value $D_{v^*} = |D_{\rho_*}| = |\rho \cos \theta|$. The second derivative of $\Psi(\pm, -)$ is $\partial^2 \Psi / \partial v^2 = \pm D_{v^*}^{-1}$.

We can evaluate the summation in Bauer's identity using the stationary phase technique. Expanding $\Psi(\pm, -)$ in a Taylor series about the stationary phase point $v^* = \rho_*$ through the quadratic term yields

$$\left. \begin{aligned} \Psi(\pm, -) &\doteq D_{v^*} \pm \frac{1}{2D_{v^*}} (v - v^*)^2 \mp \frac{\pi}{4}, \\ v^* &= \rho_* = \rho \sin \theta, \quad D_{v^*} = |\rho \cos \theta| \end{aligned} \right\} \quad (5.9-8)$$

Replacing the summation in Bauer's identity by an integral and using Eq. (5.9-8) we obtain

$$\left. \begin{aligned} \sum_{j=0}^{\infty} i^l (2l+1) \frac{\psi_l}{\rho} P_l(\cos \theta) &\doteq \sqrt{\frac{1}{2\pi}} \int_0^{\infty} D_v^{-1/2} \exp(\pm i \Psi(\pm, -)) dv \\ &\doteq \sqrt{\frac{\pm 1}{2\pi i D_{v^*}}} \exp(\pm i D_{v^*}) \int_0^{\infty} \exp\left(\pm i \frac{(v - \rho_*)^2}{2D_{v^*}}\right) dv = \exp(\pm i D_{v^*}) \end{aligned} \right\} \quad (5.9-9)$$

By a change of variable $\pi s^2 = (v - \rho_*)^2 / D_{v^*}$, the integral in Eq. (5.9-9) is transformed into a complex Fresnel integral. Since $\rho_* \gg 1$, this is essentially a

complete Fresnel integral equal to $(\pm 2\pi i D_{v*})^{1/2}$. Thus, Eq. (5.9-9) reduces to $\exp(\pm i D_{v*}) = \exp(i D_{\rho*})$, which equals the left side of Bauer's identity.

Note again that $2\psi_l = \xi_l^+ + \xi_l^-$. However, ξ_l^- appears in $\Psi(-, -)$, and ξ_l^+ appears in $\Psi(+, -)$. Therefore, only ξ_l^- contributes to the spectral sum in Eq. (5.9-9) for a position well within the incoming region, i.e., for an incoming wave. Only ξ_l^+ contributes for an outgoing wave. This is consistent with the asymptotic forms that ξ_l^+ and ξ_l^- assume when $\rho \gg v$. Here $\xi_l^\pm / \rho \rightarrow (\mp i)^{l+1} \exp(\pm i \rho) / \rho$. Thus, ξ_l^+ / ρ corresponds to an outgoing spherical wave and ξ_l^- / ρ corresponds to an incoming spherical wave.

5.9.5 The Electric Field for an Incoming Wave

We now evaluate the integrals in Eq. (5.9-5) for the electric field vector using the stationary phase technique. We first assume that the point (r, θ) is in the incoming region well away from a turning point, so that the negative argument asymptotic forms for the Airy functions apply. In an incoming region $\pi > \theta > \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$. If we set as a criterion for "well away" that the point (r, θ) must be such that $\hat{y} < -2$ at its stationary phase point, then by applying Bouguer's law in Eq. (5.6-3), one can show that a suitable criterion is $\theta > \pi/2 - \tilde{\alpha}(\rho_*, \rho_*) + \sqrt{2}/K_{\rho*}$. For GPS signals in the Earth's atmosphere, $\sqrt{2}/K_{\rho*} \approx 3 \text{ mrad}$. In an incoming region well away from a turning point only the phase ${}^{(i)}\Psi = \Psi(-, -)$ contributes significantly to the spectral integrals.

We denote the field in an incoming region by ${}^{(i)}E(\rho, \theta)$, and we let ${}^{(i)}\Psi = \Psi(-, -)$. Thus, ${}^{(i)}\Psi$ is the spectral density function for the phase delay at the point (ρ, θ) for an incoming wave. It follows from Eqs. (5.9-5) and (5.9-6) for $\rho > \rho_*$ that

$$\left. \begin{aligned} {}^{(i)}E_r(\rho, \theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty \left(\frac{\sin^3 \theta_v}{\cos \theta_v} \right)^{\frac{1}{2}} \exp(i {}^{(i)}\Psi(\rho, \theta, v)) dv, \\ {}^{(i)}E_\theta(\rho, \theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty (\sin \theta_v \cos \theta_v)^{\frac{1}{2}} \exp(i {}^{(i)}\Psi(\rho, \theta, v)) dv \end{aligned} \right\} \quad (5.9-10)$$

To evaluate these integrals using the stationary phase technique, we expand ${}^{(i)}\Psi$ in a Taylor's series through second order in spectral number about its stationary phase value. Using Eqs. (5.9-2) and (5.9-6), $\partial^{(i)}\Psi / \partial v$ becomes

$$\frac{\partial}{\partial v}({}^{(i)}\Psi) = -\left(\theta_v + \theta - \pi + \frac{\partial G}{\partial v}\right) = 0 \quad (5.9-11)$$

and the second derivative becomes

$$\frac{\partial^2}{\partial v^2}({}^{(i)}\Psi) = -\frac{1}{D_v} - \frac{\partial^2 G}{\partial v^2} \quad (5.9-12)$$

Let v^* be a spectral number value that provides a stationary phase, that is, it that satisfies the stationarity condition in Eq. (5.9-11). Then from Eqs. (5.9-2) and (5.9-11) one obtains

$$\theta_{v^*} = \pi - \theta - \left(\frac{\partial G[\rho, v]}{\partial v}\right)_{v=v^*}, \quad v^* = \rho \sin \theta_{v^*} \quad (5.9-13)$$

We define $\delta(\rho, \theta, v)$ by

$$\theta_v = \pi - \theta - \delta, \quad \delta = \pi - \theta - \sin^{-1}\left(\frac{v}{\rho}\right), \quad \rho = kr n(\rho) \quad (5.9-14)$$

where for a thin atmosphere $\delta^* = \delta(\rho, \theta, v^*)$ will generally be a small quantity. When v assumes a value so that $\delta^* = \partial G / \partial v$, Eq. (5.9-14) shows that we have a stationary phase point. In our discussion in Section 3.12 on stationary phase processes in Mie scattering theory, we noted that δ^* should be very close in value to the refractive bending angle. Here δ^* should be close in value to the cumulative bending angle $\tilde{\alpha}(\rho, \rho_*)$ for a ray passing through the point (ρ, θ) with an impact parameter value of ρ_* . Here $\rho_* = \rho \sin[\theta + \tilde{\alpha}(\rho, \rho_*)]$. From Eq. (5.7-8) it follows that $\partial G[\rho, v] / \partial v = \tilde{\alpha}(\rho, v)$ to high accuracy provided that $v < \rho - \sim 2K_v$ (see Figure 5-4). The difference between δ^* and $\tilde{\alpha}(\rho, \rho_*)$, and the between v^* and ρ_* , can be obtained by expanding Eq. (5.9-14) in a power series for $G[\rho, v]$ about the point \tilde{v} where $\partial G / \partial v = \tilde{\alpha}$, exactly. One obtains

$$\left. \begin{aligned} v^* - \rho_* &\doteq (\delta^* - \tilde{\alpha})\rho \cos(\theta + \tilde{\alpha}), \\ \delta^* - \tilde{\alpha} &\doteq (v^* - \tilde{v})\left(\frac{\partial^2 G[\rho, v]}{\partial v^2}\right)_{v=v^*} \end{aligned} \right\} \quad (5.9-15)$$

We expect that \tilde{v} will be close to ρ_* . Upon setting $\partial^2 G / \partial v^2 = \partial \tilde{\alpha} / \partial \rho_*$, it follows that

$$\left. \begin{aligned} v^* - \rho_* &= D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \left(1 - D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \right)^{-1} \\ v^* - \tilde{v} &= (\rho_* - \tilde{v}) \left(1 - D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \right)^{-1} \end{aligned} \right\} \quad (5.9-16)$$

It follows that

$$\delta^* - \tilde{\alpha} \doteq (\rho_* - \tilde{v}) \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \left(1 - D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho, \rho_*)}{\partial \rho_*} \right)^{-1} \quad (5.9-17)$$

We note the defocusing factor showing up in these expressions for the offsets. It follows from Eqs. (5.9-14) and (5.9-15) that if v^* is near the value ρ_* , which is the value(s) of the impact parameter for the ray(s) passing through the point (r, θ) , then δ^* will equal the corresponding value of $\tilde{\alpha}$ with close accuracy. If the ray system from the actual refractivity profile generates a caustic surface, when the position (r, θ) is such that v^* lies near a contact point with that caustic, $1 - (\partial \tilde{\alpha} / \partial \rho_*) D_{\rho_*} \approx 0$. Eqs. (5.9-16) and (5.9-17) are not valid in that neighborhood.

Figure 5-17 shows the stationary phase solution for the exponential refractivity profile in Eq. (5.8-2a) at the intersection of the $\partial G / \partial v$ and $\delta(\rho, \theta, v)$ curves (see Eqs. (5.7-7) and (5.9-14)). The figure shows the stationary phase solution for 4 position points (r, θ) , all at the same radial distance r_o and with θ increasing from the turning point at $\theta = \theta_* = \pi/2 - \tilde{\alpha}(\rho_o, \rho_o)$ in 4 equally spaced increments up to $\theta = \pi/2 + 3\tilde{\alpha}(\rho_o, \rho_o)$. Even for $\theta = \pi/2$, which is very close to the turning point where one might worry that either, a) the asymptotic form at negative arguments for the Airy functions might break down, or b) the approximations for the spectral coefficients given in Eq. (5.5-21) for a_l^- and Eq. (5.8-24) for a_l^+ might fail. At $\theta = \pi/2$ the intersection point yields a stationary phase value that corresponds to an Airy function argument of $\hat{y}^* \approx -20$. This value for \hat{y}^* is also corroborated by evaluating $K_\rho^{-1}(\rho_* - \rho)$ from Bouguer's law in Eq. (5.6-3). This justifies our use of the negative argument asymptotic forms for the Airy functions in the expressions for $\Psi(-, -)$ in Eq. (5.9-1) when θ is at least $\pi/2$ or greater. The intersection point for $\theta = \pi/2$ in Figure 5-17 yields a value for $\partial G / \partial v$ of 8.056 mrad, which differs by only $\sim 5 \mu\text{rad}$ from the value for $\tilde{\alpha}(\rho_o, \rho_*)$, $\rho_* = \rho_o \cos \Delta$, predicted by ray theory in Eq. (5.6-2). A lower bound on θ has already been established from Bouguer's law where we

set $\hat{y}^* = -2$; this gave $\theta \geq \pi/2 - \tilde{\alpha}(\rho_o, \rho_o) + \varepsilon$, $\varepsilon = \sqrt{2K_{u_o}^{-1}} \approx \tilde{\alpha}/3$. Thus, except for a very narrow range of θ values near the turning point, the stationary phase point v^* should be very close to ρ_* when the curvature in the refractivity is relatively slight, i.e., when $K/(kH)$ is small. $K/(kH)$ essentially is the ratio of the Airy function transition scale (a change in ρ and/or v corresponding to $\Delta\hat{y}=1$) to the refractivity scale height (in phase units), which for the example shown in Figure 5-17 is about 0.002.

Even at the turning point at $\theta = \theta_* = \pi/2 - \tilde{\alpha}(\rho_o, \rho_o)$, the intersection point in Figure 5-17 (where $\hat{y} \cong 0$ in this case) yields a value $\delta(\rho, \theta, v^*) = \partial G / \partial v|_{v^*} = 9.894 \text{ mrad}$; the ray theory value from Eq. (5.6-2) for $\tilde{\alpha}(\rho_o, \rho_o)$ is 10.110 mrad. We have already noted the levels of agreement between these quantities, which are shown in detail in Figures 5-5 and 5-6. These figures show the close agreement between the wave-theoretic and geometric optics representations of bending angle under thin atmosphere conditions except in the immediate vicinity of a turning point.

5.9.6 Evaluating the Electric Field Vector Using the Stationary Phase Technique

We now evaluate the integrals in Eq. (5.9-10) for the electric field vector in the incoming region using the stationary phase technique. We assume that the point (r, θ) is well into the incoming region so that the negative argument asymptotic forms for the Airy functions apply. We have just seen that the criterion $\theta > \pi/2 - 2\tilde{\alpha}(\rho_*, \rho_*)/3$ should suffice. Inserting into Eq. (5.9-10) the Taylor's series expansion through second order terms for $^{(i)}\Psi(\rho, \theta, v) = \Psi(-, -)$ evaluated at the stationary phase point, and using Eqs. (5.9-11)-(5.9-15), one obtains for $E_r(r, \theta)$ and $E_\theta(r, \theta)$

$$\left. \begin{aligned}
({}^i)E_{\left\{ \begin{smallmatrix} r \\ \theta \end{smallmatrix} \right\}} &\doteq E_o e^{i({}^i)\Phi^*} \left\{ \begin{aligned} &\sin({}^i)\gamma^* \\ &\cos({}^i)\gamma^* \end{aligned} \right\} \sqrt{\frac{-\sin({}^i)\gamma^*}{2\pi i D_{v^*} \sin \theta}} \int_0^\infty e^{-i \frac{(v-v^*)^2}{2\zeta D_{v^*}}} dv, \\
({}^i)\Phi^* &= ({}^i)\Psi(\rho, \theta, v^*) - \frac{\pi}{4} = \left(-D_v + v \frac{\partial G}{\partial v} - G \right)_{v=v^*}, \\
\left(\frac{\partial({}^i)\Psi}{\partial v} \right)_{v=v^*} &= \sin^{-1} \left(\frac{v^*}{\rho} \right) - \left(\pi - \theta - \frac{\partial G}{\partial v} \right)_{v=v^*} = 0, \\
({}^i)\zeta &= \left(1 + ({}^i)D_v \frac{\partial^2 G}{\partial v^2} \right)_{v=v^*}^{-1}, \quad G = G[\rho, v], \\
({}^i)\gamma^* &= \theta + \left(\frac{\partial G}{\partial v} \right)_{v=v^*}, \quad \pi > ({}^i)\gamma^* > \pi/2, \\
({}^i)D_{v^*} &= -\rho \cos({}^i)\gamma^* \geq 0, \quad v^* = \rho \sin({}^i)\gamma^*
\end{aligned} \right\} \quad (5.9-18)$$

The integrals in Eq. (5.9-18) for $({}^i)E_r$ and for $({}^i)E_\theta$ (upon noting that $-v^* \cong -\infty$) are readily transformed through a change of variable into the form $\int_{-\infty}^\infty \exp(-i\pi t^2/2) dt$, which is the complex conjugate of the complete Fresnel integral, with a value of $1-i = (-2i)^{1/2}$. Hence, the electric field for the incoming wave in Eq. (5.9-18) may be written as

$$\left. \begin{aligned}
({}^i)E_r &\doteq E_o \sqrt{\frac{({}^i)\zeta \sin({}^i)\gamma^*}{\sin \theta}} \sin({}^i)\gamma^* \exp(i({}^i)\Phi^*) \\
({}^i)E_\theta &\doteq E_o \sqrt{\frac{({}^i)\zeta \sin({}^i)\gamma^*}{\sin \theta}} \cos({}^i)\gamma^* \exp(i({}^i)\Phi^*)
\end{aligned} \right\} \quad (5.9-19)$$

We recognize in Eqs. (5.9-18) and (5.9-19) the phase delay term $({}^i)\Phi^*$, which here is largely a phase advance for the incoming wave because it is referenced to the $\theta = \pi/2$ line. Recalling the asymptotic relationships between $G[\rho, v]$ and $\partial G / \partial v$ and bending angle quantities given by Eqs. (5.7-6) and (5.7-8), respectively, we see that the equivalent form for $({}^i)\Phi^*$ in geometric optics is

$$\left. \begin{aligned} {}^{(i)}\Phi^* &\doteq \rho \cos(\theta + \tilde{\alpha}(\rho, \rho_*)) + \rho_* \tilde{\alpha}(\rho, \rho_*) + \\ &\quad \int_{\rho_*}^{\rho} \tilde{\alpha}(\rho, \nu) d\nu + \int_{\rho}^{\infty} \tilde{\alpha}(\omega, \omega) d\omega, \\ {}^{(i)}\gamma^* &\doteq \theta + \tilde{\alpha}(\rho, \rho_*) \end{aligned} \right\} \quad (5.9-20)$$

The leading term $\rho \cos(\theta + \tilde{\alpha}(\rho, \rho_*))$ is the optical path delay (negative for this incoming case) from the tangent point on the spherical surface radius ρ_* to the point (ρ, θ) . The second term is an arc length term along a sphere of radius ρ_* due to refractive bending. The third and fourth terms (equal to $-G[\rho, \nu^*]$ in wave theory (see Eq. (5.7-6)) account for the extra phase delay due to the refractive gradient-induced bending through the atmosphere that a ray of impact parameter value ρ_* from the GPS satellite (assumed to be at infinity) undergoes in reaching the incoming point (r, θ) .

Eq. (5.9-20) may be compared to the relationship with the eikonal equation given in Eq. (5.10-11). Here the spectral density function for the complete phase delay at (r, θ) , ${}^{(i)}\Psi(\rho, \theta, \nu)$, has a close correspondence to the eikonal $\mathcal{S}(\rho)$ associated with the path delay along a ray. When ${}^{(i)}\Psi(\rho, \theta, \nu)$ is evaluated at a stationary phase point $\nu^* \doteq \rho_*$, then ${}^{(i)}\Psi(\rho, \theta, \rho_*) \Leftrightarrow \mathcal{S}(\rho) - \mathcal{S}(\rho_*)$.

We also recognize the defocusing factor ${}^{(i)}\zeta$ in Eq. (5.9-18), which has its analogue in geometric optics to account for the dispersive effect of the refraction gradient on signal power (see Section 5.6). From Eq. (5.7-7), $\partial^2 G / \partial \nu^2$ is given by

$$\frac{\partial^2 G}{\partial \nu^2} \doteq 2\pi \int_{\rho}^{\infty} \frac{d \log n}{d \rho'} (\text{Ai}[\hat{y}] \text{Ai}'[\hat{y}] + \text{Bi}[\hat{y}] \text{Bi}'[\hat{y}]) d\rho' \quad (5.9-21)$$

Using the asymptotic forms for the Airy functions given in Eq. (3.8-10) when $\rho > \nu$, one obtains

$$\frac{\partial^2 G}{\partial \nu^2} \doteq -\nu^2 \int_{\rho}^{\infty} \frac{d \log n}{d \rho'} \frac{d \rho'}{(\rho'^2 - \nu^2)^{3/2}} = \frac{\partial \tilde{\alpha}}{\partial \nu}, \quad \tilde{\alpha} = \tilde{\alpha}(\rho, \nu) \quad (5.9-22)$$

The integral here can be integrated by parts to eliminate its near-singularity. Figure 5-5 compares $\partial^2 G / \partial \nu^2$ and $\partial \tilde{\alpha} / \partial \nu$ for the exponential refractivity profile given in Eq. (5.8-2a). It follows that in terms of geometric optics quantities the defocusing factor can be written as

$${}^{(i)}\zeta^{-1} \doteq \left(1 - \rho \cos(\theta + \tilde{\alpha}(\rho, \nu)) \frac{\partial \tilde{\alpha}(\rho, \nu)}{\partial \nu} \right)_{\nu=\rho_*} \quad (5.9-23)$$

When $\partial \tilde{\alpha} / \partial \nu > 0$ then ${}^{(i)}\zeta < 1$ because $\pi > \theta + \tilde{\alpha}(\rho, \rho_*) > \pi/2$.

When the refractivity profile results in a unique stationary phase path from the GPS satellite to the point (r, θ) (e.g., Case a)), then Eq. (5.9-19) shows that the refracted wave, which was planar before entering the atmosphere, is still nearly planar (for a thin atmosphere) at the point (r, θ) . However, its Poynting vector \mathbf{S} is pointed away from the original axis of propagation by an angular displacement $\tilde{\alpha}$. Its phase is delayed by an additional amount $({}^{(i)}\Phi^* - \rho \cos \theta)$. Its amplitude is modified by the factor $({}^{(i)}\zeta \sin(\theta + \tilde{\alpha}) / \sin \theta)^{1/2}$, which may be significantly greater or less than unity depending on the magnitude of $\partial^2 G / \partial \nu^2$ at the stationary phase point.

If we have a multipath situation where, for example, three different total bending angle values satisfy the boundary conditions, then Eq. (5.9-19) will appear in triplet form (or even in higher odd multiples for more complicated refractivity profiles). The Gaussian refractivity profile given in Eq. (5.8-2b) will produce triplets for a certain range of altitudes. The total field in this case would be obtained by vector addition of the field components from each contribution. We note again the failure of the stationary phase technique when the stationary phase points are too near each other or, alternatively, when they are too near the first contact points with the caustic surfaces for the complete ray system generated by varying the position of the point (r, θ) . In wave theory the accuracy of the stationary phase technique for evaluating the integrals in Eq. (5.9-10) depends on the magnitude of $\partial^2 {}^{(i)}\Psi / \partial \nu^2$ being sufficiently large. Caustics occur when stationary phase points in spectral number space also coincide with $\partial^2 {}^{(i)}\Psi / \partial \nu^2 = 0$, or nearly so. The conditions for the validity of the stationary phase approach and the third order approach leading to the Airy function of the first kind are discussed in Appendix D. Multipath scenarios and caustics are discussed more fully in Section 5.12.

Finally, we note the breakdown in accuracy of this osculating parameter approach if one attempts to use it exactly at a turning point. As we approach a turning point, $D_{\nu_*} \rightarrow 0$. Because $\partial^2 G[\rho, \nu] / \partial \nu^2$ is finite at $\nu = \rho$, we see from Eq. (5.9-18) that the defocusing factor ζ predicted by this technique approaches unity at a turning point. But from geometric optics (Eq. (5.6-16) we know that $\zeta^{-1} \rightarrow 1 + u_* n'_* / n_*$, which results from the singularity in $\partial \tilde{\alpha}(\rho, \nu) / \partial \nu$ as the turning point is approached, i.e., $D_{\nu} \partial \tilde{\alpha} / \partial \nu \rightarrow -(dn/d\rho)(\rho/n)$ as $D_{\nu} \rightarrow 0$. Turning points using an Airy layer approach are discussed in Section 5.11

5.9.7 An Outgoing Electric Field

For a point in the outgoing region well away from a turning point, only the phase combination $^{(o)}\Psi = \Psi(+, -)$ in Eq. (5.9-6) contributes significantly to the scattering integrals for the electric field representation given in Eq. (5.9-5). For an outgoing region $0 < \theta \leq \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$, and we have

$$\left. \begin{aligned} {}^{(o)}E_r(\rho, \theta) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta}} \int_0^\infty \left(\frac{\sin^3 \theta_v}{\cos \theta_v} \right)^{\frac{1}{2}} \exp(i^{(o)}\Psi(\rho, \theta, v)) dv, \\ {}^{(o)}E_\theta(\rho, \theta) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta}} \int_0^\infty (\sin \theta_v \cos \theta_v)^{\frac{1}{2}} \exp(i^{(o)}\Psi(\rho, \theta, v)) dv \end{aligned} \right\} \quad (5.9-24)$$

Referring to Eq. (5.9-6), we see that $^{(o)}\Psi = \Psi(+, -)$ is obtained from $^{(i)}\Psi = \Psi(-, -)$ through the transformation:

$\left\{ {}^{(o)}\Psi = -^{(i)}\Psi, \quad G[\rho, v] \Rightarrow 2G[\rho^\dagger(v), v] - G[\rho, v] \right\}$. $^{(o)}\Psi$ is the spectral density function for the phase delay at the point (ρ, θ) for an outgoing wave. We need only apply this transformation to the incoming forms in Eq. (5.9-18) to obtain the stationary phase evaluation of the spectral integrals for an outgoing wave. Carrying through the stationary phase computations in Eq. (5.9-24) yields the electric field for the outgoing wave. It may be written as

$$\left. \begin{aligned} {}^{(o)}E_r &\doteq E_o \sin({}^{(o)}\gamma^*) \sqrt{\frac{{}^{(o)}\zeta \sin({}^{(o)}\gamma^*)}{\sin \theta}} \exp(i^{(o)}\Phi^*), \\ {}^{(o)}E_\theta &\doteq E_o \cos({}^{(o)}\gamma^*) \sqrt{\frac{{}^{(o)}\zeta \sin({}^{(o)}\gamma^*)}{\sin \theta}} \exp(i^{(o)}\Phi^*), \\ {}^{(o)}\gamma^* &= \theta + \left(2 \frac{dG^\dagger}{dv} - \frac{\partial G}{\partial v} \right)_{v=v^*} \end{aligned} \right\} \quad (5.9-25)$$

where the outgoing phase $^{(o)}\Phi^*$ is given by

$$\begin{aligned}
{}^{(o)}\Phi^* = & \left\{ \begin{aligned} & D_v + v \left(2 \frac{dG^\dagger}{dv} - \frac{\partial G}{\partial v} \right) - (2G^\dagger - G) \Big|_{v=v^*} \\ & \doteq D_{\rho_*} + \rho_* (2\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho, \rho_*)) + \\ & 2 \int_{\rho_*}^{\infty} \tilde{\alpha}(\omega, \omega) d\omega - \int_{\rho_*}^{\rho} \tilde{\alpha}(\rho, \omega) d\omega - \int_{\rho}^{\infty} \tilde{\alpha}(\omega, \omega) d\omega, \\ & G^\dagger = G[\rho^\dagger(v), v], \quad G = G[\rho, v], \quad D_{\rho_*} = \sqrt{\rho^2 - \rho_*^2} \end{aligned} \right\} \quad (5.9-26)
\end{aligned}$$

The leading term D_{ρ_*} is the optical path delay from the tangent point on the spherical surface of radius ρ_* to the point (ρ, θ) . The second term is an arc length term along a sphere of radius ρ_* due to cumulative refractive bending $2\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho, \rho_*)$ to the outgoing point (ρ, θ) . The first and second terms combined reference the phase to the line $\theta = \pi/2$ for an incident collimated wave originating from the direction $\theta = \pi$. The third, fourth and fifth terms in wave theory account for the extra phase delay due to the refractive gradient-induced bending through the atmosphere that a ray of impact parameter value ρ_* from the GPS satellite (assumed to be at infinity) undergoes in reaching the outgoing point (ρ, θ) .

The defocusing factor ${}^{(o)}\zeta$ in Eq. (5.9-25) also has its analogue in geometric optics to account for the dispersive effect of the refraction gradient on signal power (see Section 5.6). It is given by

$$\begin{aligned}
{}^{(o)}\zeta = & \left\{ \begin{aligned} & \left(1 - D_v \left(2 \frac{d^2 G[\rho^\dagger(v), v]}{dv^2} - \frac{\partial^2 G[\rho, v]}{\partial v^2} \right) \right) \Big|_{v=v^*}^{-1} \\ & \doteq \left(1 - D_{\rho_*} \left(2 \frac{d}{d\rho_*} (\tilde{\alpha}(\rho_*, \rho_*)) - \frac{\partial}{\partial \rho_*} (\tilde{\alpha}(\rho, \rho_*)) \right) \right)^{-1} \end{aligned} \right\} \quad (5.9-27)
\end{aligned}$$

Forms for $2d^2 G[\rho^\dagger, \rho^\dagger]/dv^2 - \partial^2 G[\rho, v]/\partial v^2$ are given in Appendix J and Eq. (5.9-22).

As has already been shown for the incoming case, Eq. (5.9-25) shows the (small) deflections in angular displacement $2\tilde{\alpha}_* - \tilde{\alpha}$. Its phase is delayed by an additional amount ${}^{(o)}\Phi^* - \rho \cos \theta$. Its amplitude is modified by the factor $(\zeta \sin(\theta + 2\tilde{\alpha}_* - \tilde{\alpha})/\sin \theta)^{1/2}$, and so on.

We have already commented in the incoming case about multipath and the possible non-uniqueness of these solutions, depending on the profile of the refractivity.

5.9.8 The Electric Field at the LEO

We assume that LEO is well out of the refracting medium and receiving a signal from an outgoing wave. The equations given for the outgoing case also describe the field at the LEO except that they are somewhat simplified because $\tilde{\alpha}(\rho, \rho_*) \rightarrow 0$ as $\rho \rightarrow \infty$. They become

$$\left. \begin{aligned} E_r(\rho_L, \theta_L) &\doteq E_o \sin \gamma^* \sqrt{\frac{\zeta \sin \gamma^*}{\sin \theta_L}} \exp(i\Phi^*), \\ E_\theta(\rho_L, \theta_L) &\doteq E_o \cos \gamma^* \sqrt{\frac{\zeta \sin \gamma^*}{\sin \theta_L}} \exp(i\Phi^*), \\ \Phi^* &= D_{v^*} + 2 \left(v \frac{dG^\dagger}{dv} - G^\dagger \right)_{v^*} \doteq D_{\rho_*} + 2\rho_* \tilde{\alpha}_* + 2 \int_{\rho_*}^{\infty} \tilde{\alpha}(v, v) dv \end{aligned} \right\} \quad (5.9-28)$$

where

$$\left\{ \begin{aligned} D_{v^*} &= \rho_L \cos \gamma^* \doteq D_{\rho_*} = \rho_L \cos(\theta + 2\tilde{\alpha}_*) \\ \zeta &= \zeta(v^*, \rho_L) = \left(1 - 2D_{v^*} \frac{d^2 G^\dagger}{dv^{*2}} \right)^{-1} \doteq \left(1 - 2D_{\rho_*} \frac{d\tilde{\alpha}_*}{d\rho_*} \right)^{-1}, \\ \gamma^* &= \theta + \left(2 \frac{dG^\dagger}{dv^*} \right) \doteq \theta + 2\tilde{\alpha}_*, \quad 0 < \gamma^* < \frac{\pi}{2}, \quad v^* = \rho \sin^{(o)} \gamma^*, \\ G^\dagger &= G[\rho^\dagger(v^*), v^*], \quad \tilde{\alpha}_* = \tilde{\alpha}(\rho_*, \rho_*), \quad \rho_* = \rho_L \sin(\theta + 2\tilde{\alpha}_*) \end{aligned} \right\} \quad (5.9-29)$$

Here it is understood that both $\tilde{\alpha}_*$ and the impact parameter ρ_* are implicit functions of the LEO position coordinates (r_L, θ_L) through the application of Bouguer's law and the total refractive bending angle given in Eq. (5.6-5).

5.10 Comparison of Geometric Optics and Wave Theory

Here we briefly review the scalar diffraction results applied to a thin phase screen model. These results are then compared with those from wave theory. Thin screen models were discussed in Chapter 2, notably in Section 2.3. A thin phase screen model provides a proxy for the actual atmospheric medium through which the electromagnetic wave travels. A thin phase screen model mimics the transmission effects of the real atmosphere on a traversing electromagnetic wave. The closeness of this match is limited, of course, by limitations in the thin screen model itself. However, it is surprising how accurate a description of certain electromagnetic processes can be achieved from the model in most situations when thin atmosphere conditions apply. Thin atmosphere conditions are defined in Chapter 2, Eqs. (2.2-8) and (2.2-9). Basically, thin atmosphere conditions imply that the curvature of the ray path is much smaller than the curvature of the ambient equipotential surface, and that the length along the ray where refracting or bending occurs is small compared to the radius of curvature of the equipotential surface. Both of these conditions are related to the refractive gradient in the atmosphere. A somewhat strengthened thin screen requirement for a perpendicular mounted thin screen is that no caustics may occur in that screen. The condition for this is given in Eq. (2.3-13). Thin phase screen models offer considerable simplification in calculations of electromagnetic processes, including refractive bending and path delay, but also of other important properties, such as diffraction, multipath, caustics and shadow zones. Chapter 2 discusses two alternate thin screen models, a planar screen oriented perpendicular to the LEO/GPS line near the point of tangency of the actual ray in the Earth's atmosphere, and an impact parameter "screen", actually the post-encounter impact parameter space curve generated by varying the impact parameter $\rho_* = ka$.

In Chapter 3 we discussed rainbow caustic effects that would be difficult, but not impossible, for a thin screen model to predict. Processes at turning points also would be difficult for a thin screen model, as well as super-refractivity situations, or in the case of the planar screen when caustics occur in the screen, i.e., when $1 + a\alpha d\alpha/da = 0$. For an exponential distributed refractivity, this is equivalent to requiring $\beta < (2\pi)^{-1/2} \approx 0.4$, where β is the ray path curvature parameter given in Eq. (2.2-9) as part of the thin atmosphere definition. This threshold $\beta = (2\pi)^{-1/2}$ is exceeded across some marine layer boundaries in the lower troposphere, but rarely exceeded at higher altitudes.

5.10.1 Comparison of Wave Theory with Geometric Optics

We now compare stationary phase terms in spectral number from wave theory with phase terms from geometric optics. In ray theory the phase at the

point (r_L, θ_L) relative to the phase at the line $\theta = \pi/2$ for a collimated incident wave traveling along the direction $\theta = 0$, is given by (see Eqs. (A-55)-(A-57))

$$\varphi = \rho_L \cos(\theta_L + \alpha_L) + \rho_* \alpha_L + \int_{\rho_*}^{\infty} \alpha_L(\omega) d\omega \quad (5.10-1)$$

In wave theory a stationary value of the spectral density function for the phase delay at the LEO, Φ_L^* , referenced to the line $\theta = \pi/2$, is given from Eq. (5.9-28) by

$$\left. \begin{aligned} \Phi_L^* &= \left(\sqrt{\rho_L^2 - v^2} + 2v \frac{dG^\dagger}{dv} - 2G^\dagger \right) \Big|_{v=v^*}, \\ \left(\theta_v - \theta_L - 2 \frac{dG^\dagger}{dv} \right) \Big|_{v=v^*} &= 0, \\ \theta_v &= \sin^{-1} \left(\frac{v}{\rho_L} \right), \quad G^\dagger(v) = G[\rho^\dagger(v), v] \end{aligned} \right\} \quad (5.10-2)$$

The second line is the stationary phase condition that the spectral number must satisfy. We have seen in Section 5.7, Eqs. (5.7-4) and (5.7-14) (see also Appendix J) that

$$\left. \begin{aligned} 2 \frac{dG^\dagger(v)}{dv} &= \alpha_L(v) + O \left[\rho \frac{d^2 n}{d\rho^2} \Big|_{\rho=v} \right], \\ 2G^\dagger(v) &\doteq - \int_v^\infty \alpha_L(\omega) d\omega \end{aligned} \right\} \quad (5.10-3)$$

This means that the term $\rho(d^2 n / d\rho^2)$ must be much smaller than α_L to maintain close agreement between spectral number in wave theory and impact parameter in ray theory. This translates into a scale height in distance (i.e., not a scale height in impact parameter) for an exponential refractivity profile that must be greater than about 1 km. This scale height limit already is super-refractive in the lower troposphere, i.e., $\beta > 1$; rays won't exist for a certain critical altitude range of tangency points lying in and below the super-refractive layer. See Section 6.4. If this stationary phase condition in Eq. (5.10-2) can be satisfied, and if the curvature condition is met, i.e., $\rho_* |d^2 n / d\rho^2|_* \ll \alpha_L(\rho_*)$, then $v^* \doteq \rho_L \sin(\theta_L + \alpha_L(v^*))$, and it follows that the stationary value for the spectral density function is given by

$$\Phi_L^* \doteq \rho_L \cos(\theta_L + \alpha_L(v^*)) + v^* \alpha_L(v^*) + \int_{v^*}^{\infty} \alpha_L(\omega) d\omega \quad (5.10-4)$$

The spectral number v^* will be close to ρ_* if the inequality $\rho_* |d^2 n / d\rho^2|_* \ll \alpha_L(\rho_*)$ holds, then the stationary value of phase delay spectral density function in Eq. (5.10-4) is essentially identical to the phase delay or eikonal function from geometric optics in Eq. (5.10-1). Therefore, when a stationary phase value for the spectral number exists, then the two systems give stationary values for the phase delay that are essentially the same when the impact parameter in ray theory is set equal to that spectral number in wave theory.

Concerning the correspondence with the thin screen models, we have seen in Chapter 2, Eq. (2.3-12), that the phase delay evaluated at the point $(\rho_*, \theta_{\rho_*})$ on the impact parameter space curve (with $\theta_{\rho_*} = \pi/2 - \alpha_L(\rho_*)$) is given by

$$\varphi_{\rho_*} = \rho_* \alpha_L + \int_{\rho_*}^{\infty} \alpha_L(\omega) d\omega \quad (5.10-5)$$

The difference between this impact parameter phase delay and the stationary value of the spectral density function Φ_L^* is simply due to the geometric delay (see Figure 5-18) between the LEO and the point $(r_{\rho_*}, \theta_{\rho_*})$ on the impact parameter space curve, that is, $\rho_L \cos(\theta_L + \alpha_L)$.

Regarding the planar thin screen model mounted perpendicular to the LEO/GPS line, the Fresnel phase function in Chapter 2, Eq. (2.5-1) is given by

$$\Phi(h, h_{LG}) = \frac{\pi}{\lambda D} (h - h_{LG})^2 + \int_h^{\infty} \alpha_L(\omega) d\omega \quad (5.10-6)$$

where h and D are in phase units. Referring to Eq. (5.10-4), the first term $\rho_L \cos(\theta + \alpha_L)$ is the optical distance from the LEO to the tangent point on a sphere of radius ρ_* , the impact parameter of the ray passing through the point (ρ_L, θ_L) . The second term is the optical distance $\rho_* \alpha_L$ along the circumference of a sphere of radius ρ_* from $\theta = \pi/2 - \alpha_L$ to $\theta = \pi/2$. The sum of these two terms is given by

$$\left. \begin{aligned} &\rho_L \cos(\theta_L + \alpha_L) + \rho_* \alpha_L = \\ &\rho_L \cos \theta_L \sec \alpha_L - \alpha_L^3 \rho_L \sin \theta_L / 3 + \rho_L O[\alpha_L^4] \end{aligned} \right\} \quad (5.10-7)$$

Thus, through second order in α_L the first two terms in the stationary phase value for the spectral density function in Eq. (5.10-4) sum to the phase accumulation along the ray asymptote from the LEO to the line $\theta = \pi/2$.

The planar thin screen distance D was rather loosely defined as the perpendicular distance from the LEO to the thin screen, see Figure 5-18. In this Figure $h = \rho_L \sin(\theta_L + \alpha_L) \sec \alpha_L - \rho_o$ and $h_{LG} = \rho_L \sin \theta_L - \rho_o$, where ρ_o provides the reference altitude. The impact parameter space curve is defined by the point $(\rho_*, \theta_{\rho_*})$ as ρ_* varies. Here $\theta_{\rho_*} = \pi/2 - \alpha_L(\rho_*)$. If in Eq. (5.10-6) we now set $D = D_S = \rho_L \cos \theta_L$, it follows upon comparison with Eq. (5.10-4) that through second order in α_L we have matched the geometric components of the phase delay in the two systems. This is about as close agreement as we should expect because the version of the Rayleigh-Sommerfeld integral for scalar diffraction that we have used in Chapter 2 is accurate only through second order in α_L .

Next to be reconciled are the phase delays in the two systems induced by the refractivity gradient, which are the integral terms in Eqs. (5.10-4) and (5.10-6). We also have not been too specific about the orientation of the thin screen. For convenience here, we place it along the line $\theta = \pi/2$. Comparing these terms requires

$$\int_{h_S}^{\infty} \alpha_L(h'_S) dh'_S = \int_{\rho_*}^{\infty} \alpha_L(\rho'_*) d\rho'_* \quad (5.10-8)$$

This must hold for all values of h_{LG} , the altitude of the LEO/GPS line in the planar thin screen model above the reference plane, or equivalently, for all locations (r_L, θ_L) of the LEO during the occultation. This condition may or may not be feasible, depending on whether or not h is a single-valued function of ρ_* . Differentiating with respect to ρ_* yields

$$\alpha_L(h) \frac{dh}{d\rho_*} = \alpha_L(\rho_*) \quad (5.10-9)$$

If no caustics occur in the screen, then at every altitude we can equate the bending angle in the thin screen at a thin screen altitude h to the atmospheric bending angle $\alpha_L(\rho_*)$. If we place the thin screen along the line $\theta = \pi/2$ then it follows that

$$h = \rho_* \sec \alpha_L - \rho_L \sin \theta_o = \rho_* \sec \alpha_L - \rho_o \quad (5.10-10)$$

In a medium with large refractive gradients, thin screen caustics where $dh/d\rho_* = 0$ can arise. The perpendicular mounted thin screen is not suitable for these situations.

5.10.2 Duality between Systems

It follows from the previous discussions that when stationary values of the spectral density function exist with respect to spectral number, and when super-refractivity situations are avoided, then we have a kind of duality between the stationary phase processes over spectral number in wave theory and over impact parameter in geometric optics. For thin atmosphere conditions away from turning points, rainbow caustics, etc., we may establish a correspondence between wave theory and geometric optics when stationary phase values in each system are assumed. For wave theory applied to a spherical refracting medium on one hand, and for the scalar diffraction integral applied to a thin-screen proxy for this refracting medium on the other hand, Table 5-1 shows the correspondence between these systems when stationary phase values are assigned in each system. Table 5-1 shows this correspondence at a LEO position. Here $G^\dagger = G^\dagger(\nu) \equiv G[\rho^\dagger(\nu), \nu]$, and $\rho^\dagger(\nu) = \nu - \hat{y}^\dagger K_{\rho^\dagger}$.

In summary, a thin phase screen model combined with scalar diffraction theory gives results that closely match those from wave theory applied to a spherical geometry when, 1) caustics do not occur in the thin screen, $dh/d\rho_* \neq 0$, which is related to the thin atmosphere conditions cited earlier, 2) the observer is relatively far from the refraction zone, and 3) certain LEO rainbow caustic and reflection locations are avoided.

On the other hand, there are situations where wave theory and wave/optics approaches give disparate results. One example is when no stationary spectral numbers occur; this corresponds to super-refractivity in geometric optics or to transition across a shadow boundary. Another example is a caustic contact point where second order geometric optics predicts infinity for the amplitude of the ray and it errors in predicting the exact location of maximum flaring. Wave theory accurately addresses these cases. Although addressed later, one also can make a close correspondence between wave theory and geometric optics results when reflections occur.

5.10.3 Amendments to Account for Wavefront Curvature from the Finite Distance of the Emitting GPS Satellite

Almost all of the discussion in this monograph has assumed an infinite distance for the emitting GPS satellite. The wave front curvature effects from an emitting GPS satellite at a finite distance can be accommodated as follows. From geometric optics, see Appendix A, Eq. (A-55), the phase delay along a ray, now referenced to the emitting GPS satellite, is given by

$$\varphi = \sqrt{\rho_L^2 - \rho_*^2} + \sqrt{\rho_G^2 - \rho_*^2} + \rho_* \alpha_L(\rho_*) + \int_{\rho_*}^{\infty} \alpha_L(\omega) d\omega \quad (5.10-11)$$

Here it is assumed the emitter and the receiver are well out of any refracting medium. Otherwise, we would have to break the terms involving α_L into separate parts, one for δ_L , the ray path deflection angle at the LEO, and one for δ_G , the ray path deflection angle at the GPS, with $\alpha_L = \delta_L + \delta_G$; see Figure A-3.

This form for the phase delay in Eq. (5.10-11) is the eikonal function $\mathcal{S}(\mathbf{r})$ in geometric optics. The eikonal equation is $|\nabla \mathcal{S}| = n$, and a constant value for \mathcal{S} defines the geometric optics equivalent of a surface of constant phase delay of the electromagnetic wave (essentially a wavefront or cophasal surface).

The main change in wave theory for a finite GPS distance is in the asymptotic form that the incoming spectral coefficients assume for large radial distances out of the atmosphere. This difference between collimated and spherical incident waves was briefly discussed in Section 5.5, Eqs. (5.5-3a) and (5.5-3b). When one accounts for wavefront curvature, the spectral density function $\Psi(+, -)$ for an outgoing wave at the LEO (see Eq. (5.9-6)) is amended to the form (with $\theta_G \equiv \pi$, always)

$$\left. \begin{aligned} \Psi &= \sqrt{\rho_G^2 - v^2} + \sqrt{\rho_L^2 - v^2} + v(\theta_v^G + \theta_v^L - \theta_L) - 2G^\dagger(v) - \frac{\pi}{4} \\ \theta_v^G &= \sin^{-1}\left(\frac{v}{\rho_G}\right), \quad \theta_v^L = \sin^{-1}\left(\frac{v}{\rho_L}\right) \end{aligned} \right\} \quad (5.10-12)$$

Then the stationary value of $\Psi(+, -)$, if one exists, is obtained by setting $\partial\Psi(+, -)/\partial v = 0$. This yields

$$\frac{\partial\Psi}{\partial v} = (\theta_v^G + \theta_v^L - \theta_L) - 2\frac{dG^\dagger}{dv} = 0 \quad (5.10-13)$$

Letting $\Phi_L^* = \Psi(+, -)|_{v=v^*} + \pi/4$ and using Eq. (5.10-3), we obtain

$$\left. \begin{aligned} \Phi_L^* &= \sqrt{\rho_G^2 - v^{*2}} + \sqrt{\rho_L^2 - v^{*2}} + 2\left(v\frac{dG^\dagger}{dv}\right)_{v^*} - 2G^\dagger(v^*) \\ &\rightarrow \sqrt{\rho_G^2 - \rho_*^2} + \sqrt{\rho_L^2 - \rho_*^2} + \rho_*\alpha_L(\rho_*) + \int_{\rho_*}^{\infty} \alpha_L(\omega)d\omega \end{aligned} \right\} \quad (5.10-14)$$

The lower line is obtained by letting $v^* \rightarrow \rho_*$, and it is in agreement with the eikonal form in Eq. (5.10-11) given from geometric optics.

5.11 The Electric Field at a Turning Point

At a turning point $r = r_*$ and $\theta = \theta_*$. We have discussed in Section 5.8 the breakdown in accuracy of the representation of the field by osculating parameters when the field is evaluated in the immediate neighborhood of a turning point. When certain scale factors permit, we can accurately approximate the medium by a spherical Airy layer in the neighborhood of a turning point. We have seen in Sections 4.7 and 4.8 for a Cartesian stratified medium, that the osculating parameter approach works well at a turning point when the Airy functions are used as basis functions. At the boundary $r = r_A$ between the Airy layer below and the medium above, Section 5.8 uses the continuity conditions from electrodynamics to match the osculating parameter solution for $r \geq r_A$ to the Airy solution for $r \leq r_A$.

The Airy layer approximation works well when the effect of curvature in the refractivity profile is sufficiently small over the width of the layer. Two key scale factors can permit this approximation. The first is related to n''/n' . For an exponential medium this factor becomes H^{-1} , where H is the scale height. In this case we want the width of the Airy layer Δr to be small enough so that in the power series expansion of the exponential representation of the medium, the quadratic and higher order terms in $\Delta r/H$ are negligible, i.e., $\Delta r/H \ll 1$. On the other hand, we don't want the boundary $r = r_A$ of the Airy layer to be chosen so close to the turning point radius that the accuracy of the osculating parameter approach has begun to deteriorate, which occurs when $\hat{y} > \sim -2$. Using Eq. (5.4-3) to express $\Delta \hat{y}$ in terms of Δr over the width of the Airy layer, we have

$$\Delta \hat{y} = -\frac{kH}{K_{\rho_A}} \frac{\Delta r}{H} \quad (5.11-1)$$

Thus, the two key factors permitting an Airy layer approximation are kH and $K_{\rho_A} = (kn_A r_A / 2)^{1/3}$; the latter is related to r_A / λ . When H corresponds to the Earth's dry air conditions and for GPS wavelengths, $kH / K_{\rho_A} \approx 500$. In this case we could set $\Delta r / H \sim 0.01$, achieving better than 0.01% accuracy in the representation of the refractivity by an Airy layer, with a boundary above which the osculating parameter approach also is sufficiently accurate. We also want to keep $\Delta \hat{y}$ small enough so that phase run-off in the Airy function approximate solution given Eq. (5.8-6) compared to the exact solution is negligible. This also is related to the size of n''/n' .

In matching the solution in the Airy layer involving the Airy function of the first kind with the osculating parameters/spherical Hankel functions that hold above the layer, we apply the continuity conditions from electrodynamics. This

resulted in the matching coefficients c_l and d_l given in Eq. (5.8-8). We may multiply these coefficients by any complex factor that we wish (for example, $i^l(2l+1)/l(l+1)$), provided that we apply that factor to both coefficients, or equivalently, to both the incoming and outgoing components of the field. Also, at the boundary the outgoing phase term $2G[\rho^\dagger, \nu] - G[\rho_A, \nu]$, which is given in Eq. (5.8-24), must equal the exact value of the round-trip phase delay through the Airy layer. This fixes ρ^\dagger at the Airy layer boundary per Eq. (5.8-15). For a given value for l , the asymptotically exact value for the round-trip phase delay through an Airy layer is given by Eqs. (5.8-12) and (5.8-13). Fortunately, we need not explicitly apply this condition; we need only the representation by the Airy function of the first kind at the turning point. But in practice the numerical integration of these wave equations has a stability problem below the turning point. The slightest phase error in the initial conditions at the boundary $r = r_A$ results in a blow-up below the turning point, or equivalently, to a leakage of the Airy function of the second kind into the numerical solution. See the discussion of the WKB solutions and Eq. (5.7-26).

Near a turning point we start with the integral version of the spectral representation for the field given in Eq. (5.8-1b). Because we are evaluating the field near a turning point we do not have unrestricted use of the negative argument asymptotic form for the Airy function. But we still can use the negative exponential form for the Legendre function, which is generally applicable for large spectral numbers and for $0 < \theta < \pi$.

We replace the form $a_l^- \xi_l^- + a_l^+ \xi_l^+$ in Eq. (5.8-1) with $\text{Ai}[\tilde{y}]$ times a spectral number-dependent coefficient. When the negative argument asymptotic forms hold at the boundary of the Airy layer, we have from Eq. (5.8-8) the connecting relationship between the solution below the boundary and the solutions above the boundary. This is given by

$$\left. \begin{aligned} 2(\text{Ai}[\tilde{y}])_A &= \frac{(1+\varepsilon)^{1/6}}{\sqrt{\pi K_A}} \left(\xi_l^+ e^{+i(\tilde{X}_l - \hat{X}_l)} + \xi_l^- e^{-i(\tilde{X}_l - \hat{X}_l)} \right)_A \\ \tilde{X}_l &= \frac{2}{3}(-\tilde{y}_A)^{3/2} + \frac{\pi}{4}, \quad \hat{X}_l = \frac{2}{3}(-\hat{y}_A)^{3/2} + \frac{\pi}{4} \end{aligned} \right\} \quad (5.11-2)$$

Here \tilde{y} is given by Eq. (5.8-6), and $\tilde{y} = \hat{y}|1 - \beta|^{-2/3}$. We also have seen in Section 5.8 that

$$(\tilde{X}_l - \hat{X}_l)_A = G[\rho_A, \nu] - G[\rho^\dagger, \nu] \quad (5.11-3)$$

where ρ^\dagger in this case is adjusted to force this equality to hold, which is a value very close to ν . In general, it follows from Eq. (5.8-8) that we may set

$$(a_l^- \xi_l^- + a_l^+ \xi_l^+) = C_l \frac{2\sqrt{\pi K_{\rho_A}}}{|1-\beta|^{1/6}} \text{Ai}[\tilde{y}_A] \quad (5.11-4)$$

where C_l is chosen to account for the asymptotic boundary conditions on the wave. For an initially collimated wave, we use the asymptotic boundary conditions as $r \rightarrow \infty$ for a planar incoming wave. For this case, $a_l^-(\rho_A)$ is given by Eq. (5.5-21) and $a_l^+(\rho_A)$ is given by Eq. (5.8-24). It then follows from Eqs. (5.11-2) and (5.11-3) that C_l is given by

$$C_l = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG^\dagger(v)), \quad G^\dagger(v) = G[\rho^\dagger(v), v] \quad (5.11-5)$$

We now have the replacement form at the boundary of the Airy layer for an originally collimated incoming wave. It is given by

$$(a_l^- \xi_l^- + a_l^+ \xi_l^+) = \frac{2\sqrt{\pi K_{\rho_A}}}{|1-\beta|^{1/6}} i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG^\dagger(v)) (\text{Ai}[\tilde{y}])_A \quad (5.11-6)$$

We can use this form on the RHS to evaluate the field at any point within the Airy layer. For a given spectral number and radial position \tilde{y} is obtained from Eq. (5.8-6). Upon using the spectral form in Eq. (5.8-1b), we obtain

$$\left. \begin{aligned} E_r(\tilde{\rho}, \theta) &= \frac{|1-\beta|^{-1/6}}{K_{\rho_*}} \frac{E_o}{\sqrt{\sin \theta}} \int_0^\infty \left(\frac{v}{\rho}\right)^{\frac{3}{2}} \text{Ai}[\tilde{y}] e^{i\left(v\left(\frac{\pi}{2}-\theta-G^\dagger(v)\right)\right)} dv \\ E_\theta(\tilde{\rho}, \theta) &= \frac{|1-\beta|^{-1/6}}{K_{\rho_*}^2} \frac{iE_o}{\sqrt{\sin \theta}} \int_0^\infty \int_0^\infty \left(\frac{v}{\rho}\right)^{\frac{1}{2}} \text{Ai}[\tilde{y}] e^{i\left(v\left(\frac{\pi}{2}-\theta-G^\dagger(v)\right)\right)} dv \\ \tilde{y} &= -\frac{1}{K_{\rho_*} |1-\beta|^{2/3}} ((1-\beta)(\tilde{\rho}-\rho_*)-(l-\rho_*)), \\ \tilde{\rho} &= kn_* r, \quad \beta = -\left(\frac{1}{n} \frac{dn}{d\tilde{\rho}} \tilde{\rho}\right)_{\rho_*}, \quad K_{\rho_*} = \left(\frac{\rho_*}{2}\right)^{\frac{1}{3}}, \quad \tilde{\rho} \leq \tilde{\rho}_A \end{aligned} \right\} \quad (5.11-7)$$

The constant gradient of the refractivity in the Airy layer, $n' = dn/d\tilde{\rho}$, is set by the actual values of $n(\rho)$ at the turning point r_* and at the boundary at $r = r_A$. Thus, the gradient of the refractivity will be discontinuous at the boundary, but refractivity itself will be continuous.

5.11.1 Fourier Transform of the Airy Function

We have seen in previous sections on stationary phase processes, that near a turning point and at the stationary phase value in spectral number, $dG[\rho^\dagger(\rho_*), \rho_*] / d\rho_* \doteq \tilde{\alpha}(\rho_*, \rho_*)$ with very high accuracy for thin atmosphere conditions. But for now let us consider the integrals in Eq. (5.11-7) as Fourier transforms. When throughout the medium $n(\rho) = \text{constant}$, then $G[\rho, \nu] \equiv 0$, or at most a constant; it can be removed from the integrands in Eq. (5.11-7). From [8] we have the integral representation for the Airy function

$$\text{Ai}[\hat{y}] = \frac{(3a)^{1/3}}{2\pi} \int_{-\infty}^{\infty} \exp\left(i(at^3 + (3a)^{1/3} \hat{y}t)\right) dt \quad (5.11-8)$$

where a is any positive constant. It is easy to show that this integral form satisfies the differential equation for the Airy function $\text{Ai}'' = \hat{y} \text{Ai}$. Using the Dirac Delta function

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega t) dt \quad (5.11-9)$$

it follows from Eq. (5.11-8) and (5.11-9) that the Fourier transform of $\text{Ai}[\hat{y}]$ is given by

$$\int_{-\infty}^{\infty} \text{Ai}[\hat{y}] \exp(i\omega \hat{y}) d\hat{y} = \exp(-i\omega^3 / 3) \quad (5.11-10)$$

If we set $\omega = K_\rho(\pi/2 - \theta)$ and we use the approximation $\nu \doteq \rho + K_\rho \hat{y} + \hat{y}^2 / 60K_\rho$, (see Eq. (5.4-3)), then we obtain

$$\int_{-\infty}^{\infty} \text{Ai}[\hat{y}] e^{i\nu(\pi/2 - \theta)} d\hat{y} = e^{i\rho(\pi/2 - \theta - (\pi/2 - \theta)^3/3! + (\pi/2 - \theta)^5/5!)} \doteq e^{i\rho \cos \theta} \quad (5.11-11)$$

Thus the Fourier transform of the Airy function of the first kind, at least near a turning point, is the phasor associated with the optical path length along the ray measured from the turning point, hardly a surprising result.

We can carry the Fourier transform approach a bit further and apply it to the radial component of the electric field in Eq. (5.11-1) for the case of the homogeneous medium where $G[\rho, \nu] \equiv 0$. Using again the Dirac delta function and its derivatives, it can be shown upon expanding the term $(l/\rho)^{3/2}$ in powers of \hat{y} that

$$\frac{1}{\sqrt{\sin \theta}} \int_{-\infty}^{\infty} \left(\frac{l}{\rho} \right)^{\frac{3}{2}} \text{Ai}(\hat{y}) e^{i\nu(\pi/2-\theta)} d\hat{y} \doteq \sin \theta e^{i\rho \cos \theta} \quad (5.11-12)$$

which we already know from physical considerations to be true.

Returning to the stratified case in Eq. (5.11-7), we now evaluate the radial component of the field at the turning point $E_r(\rho_*, \theta_*)$, where $\theta_* = \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$. We expand $G[\rho^\dagger, \nu]$ in powers of spectral number about the given value ρ_* where $(dG[\rho^\dagger(\nu), \nu]/d\nu)_{\rho_*} \doteq \tilde{\alpha}(\rho_*, \rho_*)$. We also make a change of variable in the integration from ν to \tilde{y} using Eq. (5.11-7). We obtain

$$\left. \begin{aligned} E_r(\rho_*, \theta_*) &\doteq E_o \sqrt{\frac{1-\beta}{\cos \tilde{\alpha}_*}} e^{i\Phi_{\rho_*}} \int_{-\infty}^{\infty} \text{Ai}[\tilde{y}] \exp\left(-\frac{i}{2} \frac{d\tilde{\alpha}_*}{d\rho_*} (\nu - \rho_*)^2\right) d\tilde{y} \\ \Phi_{\rho_*} &= \rho_* \tilde{\alpha}_* - G[\rho^\dagger(\rho_*), \rho_*], \quad \rho_* = k r_* n(\rho_*), \\ \theta_* &= \frac{\pi}{2} - \tilde{\alpha}_*, \quad \frac{d\tilde{\alpha}}{d\rho} = \frac{d\tilde{\alpha}(\rho, \rho)}{d\rho}, \quad \tilde{\alpha}_* = \tilde{\alpha}(\rho_*, \rho_*), \\ \nu &\doteq \rho_* + \tilde{y}(1-\beta)^{2/3} K_{\rho_*}, \quad \beta = -\left(\frac{1}{n} \frac{dn}{d\rho} \tilde{\rho}\right)_{\rho_*}, \quad \tilde{\rho} = k n_* r \end{aligned} \right\} \quad (5.11-13)$$

We note the phase term Φ_{ρ_*} . It may be written as

$$\Phi_{\rho_*} = \rho_* \tilde{\alpha}(\rho_*, \rho_*) - G[\rho^\dagger(\rho_*), \rho_*] = \rho_* \tilde{\alpha}(\rho_*, \rho_*) + \int_{\rho_*}^{\infty} \tilde{\alpha}(\omega, \omega) d\omega \quad (5.11-14)$$

The first term is simply an optical path length along the circular arc of radius ρ_* measured from the reference point ($\theta = \pi/2$) to the angular position $\theta = \pi/2 - \tilde{\alpha}_*$. The second term, from Eq. (5.7-6), is the phase retardation induced by the gradient in the refractivity that the ray incurs in traveling through the atmosphere to the tangency point at r_* .

The quadratic term in the integral in Eq. (5.11-13) involving $d^2G/d\nu^2 = d\tilde{\alpha}/d\nu$ is related to the defocusing of the incident collimated wave resulting from the gradient of the refractivity.

5.11.2 Fresnel Transform of the Airy Function

To obtain the Fresnel transform of the Airy function, we again use the integral form for the Airy function given by Eq. (5.11-8). We obtain

$$\int_{-\infty}^{\infty} \text{Ai}[\tilde{y}] \exp(i(\sigma\tilde{y} + \tau\tilde{y}^2)) d\tilde{y} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iz^3/3} \left(\int_{-\infty}^{\infty} e^{i((z+\sigma)\tilde{y} + \tau\tilde{y}^2)} d\tilde{y} \right) dz \quad (5.11-15)$$

Here σ and τ are constants. From Eq. (5.11-13) it follows for our case that

$$\tau = -\frac{1}{2} K_{\rho_*}^2 (1-\beta)^{\frac{4}{3}} \frac{d\tilde{\alpha}}{d\rho_*} \quad (5.11-16)$$

Eq. (5.11-15) is a more general version of the Fresnel transform; the latter is obtained by setting $\sigma = 0$. Upon completing the square in the inner integral of Eq. (5.11-15), one can write it in terms of the complete Fresnel integral. Then with a change of variable and the use again of the integral form for the Airy function given in Eq. (5.11-8), one finally obtains

$$\int_{-\infty}^{\infty} \text{Ai}[\tilde{y}] \exp(i(\sigma\tilde{y} + \tau\tilde{y}^2)) d\tilde{y} = \sqrt{\frac{\pi}{|\tau|}} \text{Ai}[b] \exp(ic) \quad (5.11-17)$$

where

$$\begin{cases} b = -\frac{1+8\sigma\tau}{16\tau^2} \\ c = \pm \frac{\pi}{4} + \frac{1}{192\tau^3} - \frac{(1+4\sigma\tau)^2}{64\tau^3} \end{cases} \quad (5.11-18)$$

Here $+\pi/4$ is used in the expression for c when $\tau > 0$, and $-\pi/4$ when $\tau < 0$. Thus, the Fresnel transform of the Airy function yields an Airy function again but with offsets in phase and argument and with a modified amplitude. We note the defocusing term $\sqrt{|\tau|}$ appearing in the denominator of Eq. (5.11-17).

Using the Fresnel transform of the Airy function, the value of the integral in Eq. (5.11-13) becomes

$$E_r(\rho_*, \theta_*) = \frac{E_o |1-\beta|^{1/2}}{\sqrt{\cos \tilde{\alpha}_*}} \sqrt{\frac{\pi}{|\tau|}} \text{Ai}[b] \exp(i(\Phi_{\rho_*} + c)) \quad (5.11-19)$$

We note that for thin atmosphere conditions $\tau \ll 1$. For an exponential atmosphere,

$$\tau = -\frac{1}{4} \left(\frac{\pi K_{\rho_*}}{kH} \right)^{\frac{1}{2}} \beta (1-\beta)^{\frac{4}{3}} \quad (5.11-20)$$

For GPS signals with dry air at sea level, $\tau \sim 10^{-3}$. In this case Eq. (5.11-17) essentially collapses to unity as $\tau \rightarrow 0$ and for $\sigma = 0$. In this case Eq. (5.11-19) reduces to the form

$$E_r(\rho_*, \theta_*) = \frac{E_o |1 - \beta|^{1/2}}{\sqrt{\cos \tilde{\alpha}_*}} \exp(i\Phi_{\rho_*}) \quad (5.11-21)$$

We see the defocusing factor $|1 - \beta|^{1/2}$ appearing here. Squaring this factor yields the defocusing in signal power, which essentially agrees with Eq. (5.6-16), the prediction from geometric optics for the defocusing at a turning point.

For E_θ one can use the Fresnel transform of $\text{Ai}'[\hat{y}]$ to derive similar results. E_θ should be essentially zero at a turning point. Also, the integral of $\text{Ai}'[\hat{y}]$ over the real axis is zero.

5.12 Caustics and Multipath

Whether or not there is a unique ray, or possibly no ray, arriving at the point (r, θ) depends on the nature of the profile of $n(\rho)$, which may or may not result in caustics for the ray system and consequent multiple ray paths within the multipath zone. We have discussed caustics and their effects Chapters 2 and 3, and they arise again here. These chapters also show examples of the converse of a multipath zone: the shadow zone, which in second order geometric optics is a zone devoid of rays in the limit and extreme defocusing in actuality. A shadow zone based on geometric optics has an analogue in wave theory. In shadow zones, wave theory using the spectral integrals in Eqs. (5.8-1) or (5.9-24) predicts much-diminished amplitudes for the field, but not zero. Also, wave theory accounts for diffraction. Using the stationary phase technique, it is readily shown that in a strict shadow zone in geometric optics, the wave theory spectral integrals have no stationary phase points in spectral number; thus, a small but non-zero fluctuating amplitude results. In the thin screen model geometric optics predicts no rays within a shadow zone, but the scalar diffraction integral from a thin screen yields a small but non-zero amplitude in this zone. The scalar diffraction integral applied to a thin screen integrates the path delay phasor over impact parameter space in the thin screen. Section 5.10 discussed the close correspondence between stationary phase values of spectral number in wave theory and stationary phase values of impact parameter in the scalar diffraction/thin screen. Under most conditions the scalar diffraction/thin screen model generates essentially the same values for the electromagnetic field at a point well away from the screen, including diffraction effects, that are obtained from wave theory using the spectral technique described here.

We have already shown a specific example without caustics in Figure 5-10 for a monotonic profile for the refractivity, the exponential profile given in Eq. (5.8-2a). In that example, for a given position (r_o, θ) , the stationary phase condition on wavenumber in Eq. (5.9-11) is satisfied uniquely at the intersection of the δ and $\partial G / \partial \nu$ curves versus ν . These curves are generated from Eqs. (5.9-14) and (5.7-7) respectively, for several values of θ , and they apply to an incoming region. The curve for $\partial G / \partial \nu$ in Figure 5-10 (and implicitly the δ curve) uses a refraction profile that corresponds closely to the physical conditions for the dry air component of the Earth's atmosphere with r_o corresponding to sea level. The value of ν at the intersection point of the $\partial G / \partial \nu$ and δ curves is essentially the impact parameter at the Earth's surface. Fixing (r_o, θ) to a different pair of values would cause a displacement of the δ curve through Eq. (5.9-14) and the $\partial G / \partial \nu$ curve through Eq. (5.7-7) and, therefore, to an intersection at a different bending angle $\tilde{\alpha}(\rho_o, \nu)$ and a different ν value.

5.12.1 A Numerical Example of Multipath and Caustics

As a simple first example we consider the refractivity profile from Eq. (5.8-2b), where $n(\rho)$ has a non-super-refracting Gaussian distribution. A more realistic case, to be considered later, embeds this distribution in a background distribution corresponding to the near-exponential form for dry air. This dry air distribution causes significant defocusing, which has the effect of greatly compressing the bending angle spread from the multiple rays.

In the Gaussian case $d\tilde{\alpha} / d\rho$ is not monotonic and, therefore, caustics and multipath arise for certain positions (r, θ) . We use the refractivity profile given by Eq. (5.8-2b) to obtain the electric field at the LEO. When this Gaussian profile is used in Eq. (5.6-5) a bipolar bending angle profile results, $\tilde{\alpha}(\rho_*, \rho_*)$ (see Appendix E). The total refractive bending angle at the LEO is $\alpha_L(\rho_*) = 2\tilde{\alpha}(\rho_*, \rho_*)$.

Now, we use this Gaussian profile for $n(\rho)$ in the integral expression for $2dG^\dagger(\nu) / d\nu = 2dG[\rho^\dagger(\nu), \nu] / d\nu$ given by Eqs. (5.7-7) and (5.7-13), and we generate the curve in Figure 5-19 by varying ν . In this example a particular set of values, $N_w = 0.0001$, $H_w = 1.6k$, and $\rho_w = \rho_o$, has been used. H_w is the $1-\sigma$ width of the Gaussian refractivity distribution. Here, $\rho_o \approx 2 \times 10^8$, which corresponds to near-sea level for a GPS wavelength. The peak refractivity is N_w is about 40% of the refractivity from dry air there.

We know from our earlier discussion that this $2dG^\dagger(\nu) / d\nu$ curve coincides with $2\tilde{\alpha}(\rho_*, \rho_*)$ to very high accuracy. From the stationarity

condition on $\Psi(+,-)$ (see Eq. (5.9-6)) for an outgoing wave, we have, upon setting $\partial\Psi/\partial v = 0$, the relationships

$$\delta = \theta_v - \theta_L = \sin^{-1}\left(\frac{v}{\rho_L}\right) - \theta_L, \quad \delta^* = \theta_{v^*} - \theta_L = 2\left(\frac{dG^\dagger(v)}{dv}\right)_{v^*} \quad (5.12-1)$$

where v^* denotes a stationary value of the spectral number. By varying v while fixing $\rho_L = kr_L = 1.1kr_o$ and θ_L to specific orbital values, we generate the δ curves in Figure 5-19 for three specific orbital values of θ_L . The intersections of the δ and $2dG[\rho^\dagger(v), v]/dv$ curves correspond to the δ^* and v^* values given in Eq. (5.12-1), which are stationary phase values for $\Psi(+,-)$, i.e., $\partial\Psi/\partial v = 0$ at these points. These two particular δ curves tangent to the $2dG^\dagger/dv$ curve bound the range of θ_L values where the effects of multipath are evident. In this example that range is about 17 mrad between the upper and lower caustic contacts for a typical LEO orbit. An intermediate value for θ_L yields a δ curve with three intersections with the $2dG^\dagger/dv$ curve and, therefore, three stationary phase points for $\Psi(+,-)$ in v -space. Within these two bounding δ curves three rays with different bending angles from the GPS satellite will concurrently converge to the LEO. Outside of this zone there is only a single ray and its corresponding bending angle.

The upper and lower δ curves, corresponding to tangency points on the $2dG^\dagger/dv$ curve with bending angle values of 12.42 and -5.68 mrad respectively, define the first contact points with the upper and lower caustic surfaces bounding this region. The θ_L values corresponding to these upper and lower tangency points are positions for the LEO where episodes of intense flaring can occur. Each tangency point is near the point of maximum observed amplitude. At these points the condition $2d^2G/dv^2 = \partial\delta/\partial v$ must hold. From Eq. (5.12-1) we have $\partial\delta/\partial v = 1/D_v$, where $D_v = \sqrt{\rho_L^2 - v^2}$. Therefore, from Eq. (5.9-29) we have for the defocusing factor at the LEO for a tangency point is given by $\zeta^{-1} = 1 - 2D_v d^2G/dv^2 = 0$, which defines a caustic contact point. Setting $v = v^* = \rho_*$, we conclude that a condition for no caustics is given by $2d^2G^\dagger/dv^2 \doteq 2d\tilde{\alpha}/d\rho_* < 1/D_{\rho_*}$ for all ρ_* values. Therefore, the gradient of the bending angle can be positive, but not too positive, less than about 0.3 mrad/km for a GPS wavelength and $r_L \approx 1.1r_o$. We also note that in cases where the refractivity gradient is continuous, the number of multipath rays is odd, except at caustic contact point.

If these stationary phase points are adequately separated in wavenumber (see Appendix D for an “adequacy” index), then geometric optics as an approximate technique should suffice, which treats them as corresponding to

separate rays. Geometric optics can be used with Eq. (5.6-5) to generate the ray for the impact parameter value $\rho_* = \rho_*(v^\dagger)$ corresponding to each of these intersection points. The phasors for three rays can be combined as vectors, i.e., taking into account their relative phases and amplitudes, to obtain the aggregate electric field at the point (r_L, θ_L) . However, near the caustic limits, two of these points will merge and geometric optics will become inadequate for stationary phase points in spectral number that are in the neighborhood of the caustic contact point. In wave theory the second order stationary phase technique for evaluating the integrals in Eq. (5.8-1b) over spectral number also becomes inadequate for the same reason. These spectral integrals should be directly evaluated, or a third order stationary phase technique can be used, such as that described in Appendix D and discussed later in this chapter.

A caustic contact historically is associated with laser-like rays, for example, a rainbow spectral component from a water drop. Whether or not that kind of flaring in signal amplitude actually occurs at a caustic depends on the curvature of the $2dG^\dagger/dv$ curve, that is, it depends on the magnitude of $\partial^3\Psi(+, -)/\partial v^3$ at the caustic contact point $v = v^*$ where $\partial\Psi(+, -)/\partial v = \partial^2\Psi(+, -)/\partial v^2 = 0$. If the magnitude of $\partial^3\Psi(+, -)/\partial v^3$ is relatively small at the caustic point, then a broader neighborhood of spectral numbers about the stationary phase point contributes constructively to the spectral integral evaluation. This results in a larger amplitude. In other words, the spectral width about a stationary phase point across which phase coherence in the integrand is preserved depends inversely on $\partial^3\Psi(+, -)/\partial v^3$ when both $\partial\Psi(+, -)/\partial v$ and $\partial^2\Psi(+, -)/\partial v^2$ are at or near zero. We will show later that the amplitude of the flaring at a caustic contact point is proportional to $(\partial^3\Psi(+, -)/\partial v^3)^{-1/3}$. If the amplitudes of the new rays created at a caustic contact point are relatively strong, then signal flaring is to be expected even if interference is present from a pre-existing ray. If the amplitudes of the new rays are relatively weak, then the pre-existing ray will be dominant and the signal flaring, such as it is, will show up on the envelope of the resulting interference fringes. The Gaussian refractivity profile used in this section yields the latter scenario. Section 5.13 and Chapter 6 provide examples of the former scenario with very strong nascent rays.

Figure 5-20 shows the relationship between LEO orbital angle θ_L and impact parameter ρ_* from geometric optics for the same Gaussian refractivity profile and orbit model used in Figure 5-19. This impact parameter diagram is obtained directly from Bouguer's law, $\theta_L = \sin^{-1}(\rho_*/\rho_L) - 2\tilde{\alpha}(\rho_*, \rho_*)$. Once the refractivity profile is specified, $\tilde{\alpha}(\rho_*, \rho_*)$ is obtained from Eq. (5.6-5). For $-12.2 \leq \theta_L - \theta_o \leq 4.8$ a triplet of impact parameter values and bending angles simultaneously satisfy the boundary conditions provided by the position (r_L, θ_L) of the LEO. Significant multipath occurs within this zone. Caustic

contacts lie at the points where $d\theta_L/d\rho_* = 0$. These points correspond to episodes of signal flaring and they mark the boundaries of the multipath zone, in this example at $\theta_L - \theta_o = 4.8$ and at $\theta_L - \theta_o = -12.2$. Above $\theta_L - \theta_o = 4.8$, rays (a) and (b) are non-existent in a second order geometric optics context. Only the (m) ray exists. Similarly, rays (m) and (a) do not exist below $\theta_L - \theta_o = -12.2$; only (b) rays exist.

For the occulted GPS satellite lying in the orbit plane of the LEO and setting, θ_L decreases with time at a rate of roughly 1 mrad/sec. Therefore, events unfold in Figure 5-20 from right to left. Above $\theta_L - \theta_o = 4.8$, only one stationary phase value in spectral number is realized. There is a unique value for any given orbit angle above this limit. This unique sequence of stationary phase values in spectral number versus θ_L above this limit corresponds in geometric optics to the impact parameter branch (m) in the figure, which defines the tangency points or impact parameter values generated by the (m) family of stationary phase paths, the (m) rays. When $\theta_L - \theta_o$ drops below 4.8 mrad, the first contact with a caustic is made. This results in flaring of the signal and the onset of interference from the triplet of competing stationary phase points in spectral number. In geometric optics this corresponds to the existence of three competing ray systems (m), (a) and (b); a member from each family passes through the position of the LEO. See Figure 2-4 for a ray diagram. Ray (a) is the anomalous ray, and although it is a path of stationary phase, it can be shown that this path provides a local maximum in phase delay. Paths (m) and (b) provide local minima. This multipath episode continues until the lower caustic contact point is reached at $\theta_L - \theta_o = -12.2$. Below this point a single stationary phase value in spectral number resumes, which corresponds to the lower main ray (b) and to the resumption of a smooth signal.

Figure 5-21 (a) shows the signal amplitude $|E(r_L, \theta_L)|$ at the LEO versus orbit angle $\theta_L - \theta_o$ for the same models used in Figures 5-19 and 5-20. This Figure was obtained from a numerical integration of the spectral integral for E_r given in Eq. (5.8-1b), or equivalently, Eq. (5.9-10). The numerical integration was aided with the stationary phase technique to isolate stationary phase neighborhoods, thereby reducing the computational burden¹ (see Section 2.6). The spectral coefficient $a_l^+(\rho)$ used in the integration is given in Eq. (5.8-25), and $G^\dagger(\nu) = G[\rho^\dagger(\nu), \nu]$ is given from Eq. (5.7-2) using the refractivity profile given in Eq. (5.8-2b), and the parameter values used for Figures 5-19 and 5-20.

¹ Here one uses the stationary phase technique to find all spectral numbers that yield a stationary value in $\Psi(+, -)$, in this example a maximum of three. Then in Eq. (5.9-24) it is only necessary to carry out a numerical integration over a suitable neighborhood around each of these spectral numbers, taking care to properly phase-connect these separate converged integration intervals.

This Figure shows the smooth and nearly constant signal amplitude outside of the multipath zone, and it clearly shows the abrupt onset of a multipath episode at the upper caustic contact point at about $\theta_L - \theta_o \approx 4.8$, continuing until the lower contact point at $\theta_L - \theta_o \approx -12.0$. This Figure demonstrates the violent interference from a triplet set of stationary phase points that can result from a non-monotonic refractivity profile if the refractive gradient is even moderate in magnitude. Caustic flaring at the contact points is almost immediately overwhelmed by phase interference among the three full-fledged stationary phase paths. In this example, inspection of the gradients in bending angle in the middle of the multipath zone in Figure 5-20 show that none of the paths is overly defocused relative to its competitors. Defocusing is evident in the neighborhood of $\theta_L - \theta_o = 0$ because of the steep gradients in bending angle on all three rays there. Above $\theta_L - \theta_o = 5$ the principal contribution to $E(r_L, \theta_L)$ comes from the main ray (m) without interference from the (a) and (b) rays (see Figure 5-20). Below this point three mutually interfering ray paths contribute to the spectral integrals. Contact with the lower caustic occurs at $\theta_L - \theta_o \approx -12.2$. Abrupt termination of the scintillation episode follows. Below this only the (b) ray remains to contribute to $E(r_L, \theta_L)$.

Figure 5-21 (b) is a blow-up of the neighborhoods around the upper and lower caustic contact points to show the fast and slow fringes in amplitude of the field at the LEO. The high frequency fringes from interference between rays are resolved at the upper contact point; at the lower point their spacing is less than the resolution of the figure. The lower contact point shows somewhat less flaring than the upper point. Figure 5-20 shows that $d^2\theta_L / d\rho_*^2$ is somewhat larger at the lower point than at the upper point. It will be shown that $d^2\theta_L / d\rho_*^2 = \partial^3\Psi(+, -) / \partial v^3$ when $v = \rho_*$. It follows that when $d^2\theta_L / d\rho_*^2$ is larger, the near-tangency condition between the $2dG^\dagger(v) / dv$ curve and the δ curve near a caustic contact point runs over a shorter interval in spectral number. This is so on the left caustic point in (b). Therefore, a smaller contribution to the spectral integral for $E(r_L, \theta_L)$ comes from the spectral neighborhood around the lower caustic point than from the neighborhood around the caustic point. The peak amplitudes near caustic contacts are slightly offset, 0.1–0.2 mrad, from the geometric optics predictions for caustic contact (i.e., where $\zeta \rightarrow \infty$).

Figure 5-22 shows the difference in phase near the upper caustic point between the complete field at the LEO and the field from only the (m) ray. The spikes of large phase acceleration in this Figure correspond to orbit angles in Figure 5-21 (b) where the complete field is almost totally annihilated through mutual interference between the three rays. The frequency of the phase fringes scales with impact parameter separation of the rays, which in turn scales roughly linearly with the $1 - \sigma$ width of the Gaussian refractivity profile used in

these figures. Fast fringes are from interference between the main ray (m) and the nascent rays (a) and (b). Slow modulation or banding results from interference between the (a) and (b) rays as their impact parameters begin to separate near the caustic point.

5-12-2 A Geometric Optics Interpretation of Multipath

Geometric optics can accurately predict many of the fringe features for multipath situations, such as those shown in Figures 5-21 and 5-23. One must take into account the amplitudes of the competing rays and their phase interference as the orbit angle or equivalently, time, evolves. A notable exception is the failure of geometric optics to accurately predict the field in the immediate vicinity of a caustic contact. Second order geometric optics predicts an infinite amplitude at such places, whereas wave theory gives the correct values.

We now apply geometric optics to see how well it does at predicting the fringe amplitudes and frequencies shown in Figure 5-21. Starting with Bouguer's law, and given our specific refractivity and orbit models, the impact parameter curve for each competing ray system can be calculated. From the impact parameter diagram in Figure 5-20 one can determine the defocusing of any ray versus orbit angle and also the interference frequencies between any two rays. To interpret the fringes, we need the relative phase changes and amplitudes of these competing rays evaluated at the LEO.

For the phase change or excess Doppler, we start with the stationary value of the spectral density of the phase at the LEO. From Eqs. (5.9-6) and (5.12-1) it is given by

$$\left. \begin{aligned} \Psi^*(+, -) &= D_{v^*} + v^* (\theta_{v^*} - \theta_L) - 2G^\dagger(v^*) - \frac{\pi}{4}, \\ D_{v^*} &= \sqrt{\rho_L^2 - v^{*2}}, \quad v^* = \rho_L \sin \theta_{v^*} \end{aligned} \right\} \quad (5.12-2)$$

where $G^\dagger(v) = G[\rho^\dagger(v), v]$ is given in Eq. (5.7-2) with $\rho^\dagger = v - \hat{y}^\dagger K_{\rho^\dagger}$. Here we evaluate $\Psi(+, -)$ at a stationary value $\Psi^*(+, -)$ with respect to spectral number. That spectral number v^* is allowed to assume the value of the impact parameter ρ_* at a given epoch for a specific ray, either (m), (a) or (b) in Figure 5-20. We know from spectral theory that v^* will be very close to ρ_* for that ray, away from super-refracting conditions. Now we differentiate $\Psi^*(+, -)$ with respect to time. Both the orbital position of the LEO and the impact

parameter change as time evolves². Since we are interested only in the difference in phase between rays, we can neglect the Doppler term from the observed GPS satellite. Thus, we have

$$\left. \begin{aligned} \frac{d\Psi^*}{dt} &= \frac{\partial\Psi^*}{\partial\rho_L} \dot{\rho}_L + \frac{\partial\Psi^*}{\partial\theta_L} \dot{\theta}_L + \frac{\partial\Psi^*}{\partial\nu} \dot{\rho}_* \\ &= \frac{D\rho_*}{\rho_L} \dot{\rho}_L - \rho_* \dot{\theta}_L + \dot{\rho}_* \left(\theta_v - \theta_L - 2 \frac{dG^\dagger}{d\nu} \right)_{\nu^*=\rho_*} \end{aligned} \right\} \quad (5.12-3)$$

But, at a stationary phase point in spectral number, $\theta_v - \theta_L - 2dG^\dagger/d\nu = 0$, and for a circular orbit $\dot{\rho}_L \equiv 0$. It follows in this case that

$$\left. \begin{aligned} \frac{d\Psi^*}{dt} &= -\rho_* \dot{\theta}_L \quad \text{or} \quad \frac{d\Psi^*}{d\theta_L} = -\rho_* \end{aligned} \right\} \quad (5.12-4)$$

Thus, $d(\Delta\Psi_*)/d\theta_L = -\Delta\rho_*$, where $\Delta\rho_*$ is the difference in impact parameter values between any two of the three rays (m), (a) or (b) at a given epoch or orbit angle, and $\Delta\Psi_*$ is the difference in phase at the LEO between these two rays. These impact parameter differences can be read directly from Figure 5-20. The high frequency fringes in Figure 5-21 (b) near $\theta_L - \theta_o = 5$ result from interference between the combined but relatively weak³ (a) and (b) rays with the strong (m) ray. At $\theta_L - \theta_o = 4.8$, the difference in altitudes of the impact parameters between the (m) ray and the nascent (a) and (b) rays is 16.3 km. Thus the frequency of the fringes resulting from phase interference between these two ray systems is given by $(2\pi)^{-1} d(\Psi_{*a} - \Psi_{*bc})/d\theta_L = 16.3/\lambda$, or about 82 cycles/mrad. Although difficult to measure precisely in Figure 5-21 (b), the high frequency fringe rate is indeed about 80 cycles/mrad. At the lower caustic contact point near $\theta_L - \theta_o = -12$, the high frequency fringes are caused

² We also should include the obliquity effect resulting from the GPS satellites generally not lying in the LEO orbit plane. This correction factor is discussed in Chapter 6. It can reduce Doppler values by up to about 35% for inclination angles up to 30 deg. Here we assume that the occulted GPS satellite lies in the orbit plane of the LEO. This assumption essentially impacts only the scale factor between LEO orbit angle change and elapsed time.

³ Caustic rays are renown for beaming like searchlights at their nascence, but in this example the (a) and (b) rays are relatively dim (combined amplitude of 1.14 compared to the (m) ray amplitude of 1.0). This is because of the relatively large magnitude of $d^2\theta_L/d\rho_*^2$ at the caustic contact point. The upper Gaussian wing of the refractivity profile leads to a rapid falloff in the gradient of the bending angle with increasing altitude and therefore, to a main ray (m) not significantly defocused here.

by interference between the (b) ray and the dying (m) and (a) rays (or nascent rays for a rising occultation). Figure 5-20 shows a much wider separation in impact parameter values there between the (b) ray and the (m) or (a) rays. The high frequency fringe rate is around 180 cycles/mrad. We deal with fringe frequency compression from defocusing later.

The low-frequency modulation in Figure 5-21 (b) comes from interference between the (a) and (b) rays near the upper caustic contact point at $\theta_L - \theta_o = 4.8$, and between the (m) and (a) rays near the lower contact at $\theta_L - \theta_o = -12$. These ray pairs have narrowly different impact parameter values at these locations in this figure, but they have about the same amplitude initially. For the right hand side of (b) near $\theta_L - \theta_o = 4$, $(2\pi)^{-1} d(\Psi_b^* - \Psi_c^*) / d\theta_L$ is about 5 cycles/mrad, but this modulation frequency is accelerating rapidly with decreasing θ_L . We can infer this acceleration from Figure 5-20, which shows the impact parameter values of these two rays separating rapidly with decreasing θ_L as they move away from the caustic contact point where these rays were created. At the lower caustic point the low frequency and weakening modulation results from interference between the (m) and (a) rays, which are dying out completely below $\theta_L - \theta_o = -12$.

5.12.3 Amplitude Variability from Geometric Optics

Except in the immediate neighborhood of the caustic points, the amplitude variability shown in Figure 5-21 can be predicted using geometric optics. In this approach one adds up the complex amplitudes of the three rays taking into account their respective defocusing. The amplitude of the slow modulation in this Figure can be obtained from the defocusing factor $\zeta^{-1} = 1 - 2D_{\rho_*} d\tilde{\alpha}_* / d\rho_*$ for each ray. The slow modulation peaks and valleys in amplitude shown in Figure 5-21 (b) are accurately predicted from the four combinations $|\zeta_c^{1/2} \pm \zeta_a^{1/2} \pm \zeta_b^{1/2}|$. (ζ_b is negative, a hallmark of an anomalous ray; the path delay along this ray is a local maximum.) For example, in panel (b) at $\theta_L - \theta_o = 4.2$, the (m) ray is still not significantly defocused; its amplitude is unity. For the (a) and (b) rays we have $|\zeta_b|^{1/2} = 0.382$ and $\zeta_c^{1/2} = 0.461$. The four combinations yield $1 \pm 0.461 \pm 0.382 = 1.843, 0.157, 1.079$, and 0.921 . The corresponding peak values of the envelope in (b) are about 1.84 and 0.16 for the points of maximum amplitude difference, and 1.08 and 0.93 for the nodes. This is very close agreement, considering that the envelope isn't well defined in that figure.

At the upper caustic contact point in Figure 5-21 (b), $\zeta_{bc} \rightarrow \infty$; therefore, the agreement cited above must break down as we near such a point. From geometric optics this occurs at $\theta_L - \theta_o = 4.77$, but the peak amplitude from

wave theory in (b) occurs at $\theta_L - \theta_o = 4.65$. At the lower caustic point, $\zeta_{ab} \rightarrow \infty$ at $\theta_L - \theta_o = -12.18$, but the peak amplitude from wave theory occurs at $\theta_L - \theta_o = -11.98$.

5.12.4 Third Order Theory

The small offsets in location of the peak and the major discrepancy in amplitude prediction from geometric optics can be reconciled through use of a third order stationary phase technique. Here, to evaluate the spectral integrals in Eq. (5.9-10), we expand $\Psi(+, -)$ in spectral number about the caustic contact point where $\partial^2 \Psi(+, -) / \partial v^2 = 0$, retaining third order terms. We obtain

$$\Psi(+, -) \doteq \Psi|_{v_{\mathcal{C}}} + (v - v_{\mathcal{C}}) \frac{\partial \Psi}{\partial v} \Big|_{v_{\mathcal{C}}} + \frac{(v - v_{\mathcal{C}})^3}{3!} \frac{\partial^3 \Psi}{\partial v^3} \Big|_{v_{\mathcal{C}}}, \quad \frac{\partial^2 \Psi}{\partial v^2} \Big|_{v_{\mathcal{C}}} = 0 \quad (5.12-5)$$

Here $v_{\mathcal{C}}$ is the value of the spectral number at the caustic contact point. It satisfies the tangency condition between the $2dG^\dagger / dv$ and δ curves shown in Figure 5-19, and we know that its value is close in value to $\rho_{\mathcal{C}}$, which is the impact parameter value at the caustic contact point. Therefore, for the purpose of evaluating the amplitude of the field at the LEO we set $v_{\mathcal{C}} = \rho_{\mathcal{C}}$. The partial derivatives $\partial \Psi(+, -) / \partial v$ and $\partial^3 \Psi(+, -) / \partial v^3$ are evaluated at that point. We obtain

$$\left. \begin{aligned} \frac{\partial \Psi}{\partial v} \Big|_{v=\rho_{\mathcal{C}}} &\doteq \Psi'_{\mathcal{C}} = \theta_{\rho_{\mathcal{C}}} - \theta_L - 2\tilde{\alpha}_{\mathcal{C}}, \\ \frac{\partial^2 \Psi}{\partial v^2} \Big|_{v=\rho_{\mathcal{C}}} &\doteq \Psi''_{\mathcal{C}} = \frac{d\theta_L}{d\rho_*} \Big|_{\rho_*=\rho_{\mathcal{C}}} = \left(\frac{1}{D\rho_*} - 2 \frac{d\tilde{\alpha}}{d\rho_*} \right)_{\rho_{\mathcal{C}}} = 0, \\ \frac{\partial^3 \Psi}{\partial v^3} \Big|_{v=\rho_{\mathcal{C}}} &\doteq \Psi'''_{\mathcal{C}} = \frac{d^2 \theta_L}{d\rho_*^2} \Big|_{\rho_*=\rho_{\mathcal{C}}} = \left(\frac{\rho_*}{D^3 \rho_*} - 2 \frac{d^2 \tilde{\alpha}}{d\rho_*^2} \right)_{\rho_{\mathcal{C}}}, \\ \tilde{\alpha}_{\mathcal{C}} &= \tilde{\alpha}(\rho_{\mathcal{C}}, \rho_{\mathcal{C}}), \quad 2\tilde{\alpha} = \alpha_L \end{aligned} \right\} \quad (5.12-6)$$

Here the relationships, $\Psi'' = d\theta_L / d\rho_*$ and $\Psi''' = d^2 \theta_L / d\rho_*^2$, follow from Bouguer's law in Eq. (5.12-5). Differentiating that expression successively with respect $\rho_{\mathcal{C}}$ yields the relationships given in Eq. (5.12-6). Since we seek only the modulation amplitude near the caustic contact point (but not the phase), we can simplify the spectral integrals Eq. (5.9-10) by placing the slowly varying

quantities (i.e., $(\sin^3 \theta_v / \cos \theta_v)^{1/2}$ and $(\sin \theta_v \cos \theta_v)^{1/2}$) outside of the integrals and setting them equal to their values at the caustic contact point. When the radial and angular components of the field are combined, Eq. (5.9-10) yields the field at the LEO from the nascent rays near the caustic contact in the form

$$E_{\mathcal{C}}(r_L, \theta_L) \doteq E_o \left(\frac{\tan \theta_{\mathcal{C}}}{2\pi \rho_L \sin \theta_L} \right)^{\frac{1}{2}} e^{i\Psi_{v_{\mathcal{C}}}} \int_0^{\infty} e^{i \left((v-v_{\mathcal{C}}) \Psi'_{v_{\mathcal{C}}} + \frac{1}{3!} (v-v_{\mathcal{C}})^3 \Psi'''_{v_{\mathcal{C}}} \right)} dv \quad (5.12-7)$$

Here $\sin \theta_{\mathcal{C}} = \rho_{\mathcal{C}} / \rho_L$, where $\rho_{\mathcal{C}}$ is the impact parameter at the caustic contact point. The subscript “ \mathcal{C} ” on $E_{\mathcal{C}}(r_L, \theta_L)$ is to remind us that it represents the nascent field only. The Taylor series expansion in Eq. (5.12-7) is not intended to span the spectral number range needed to include the stationary point corresponding to the main ray. Making a change of integration variable to $(v - v_{\mathcal{C}}) = \text{sign}[\Psi'''_{v_{\mathcal{C}}}] \left(2 / |\Psi'''_{v_{\mathcal{C}}}| \right)^{1/3} t$, and noting that $v_{\mathcal{C}} \doteq \rho_{\mathcal{C}}$ is a very large number, we obtain for the amplitude of the nascent field at the LEO

$$\left. \begin{aligned} |E_{\mathcal{C}}(r_L, \theta_L)| &\doteq E_o \left(\frac{2\pi \tan \theta_{\mathcal{C}}}{\rho_L \sin \theta_L} \right)^{\frac{1}{2}} \left(\frac{2}{|\Psi'''_{v_{\mathcal{C}}}|} \right)^{\frac{1}{3}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i(z t + t^3 / 3)) dt = \\ E_o \left(\frac{2\pi \tan \theta_{\mathcal{C}}}{\rho_L \sin \theta_L} \right)^{\frac{1}{2}} \left(\frac{2}{|\Psi'''_{v_{\mathcal{C}}}|} \right)^{\frac{1}{3}} \text{Ai}[z] &= E_o \left(\frac{2\pi \tan \theta_{\mathcal{C}}}{\rho_L \sin \theta_L} \right)^{\frac{1}{2}} \left(\frac{1}{2} \left| \frac{d^2 \theta_L}{d\rho_{\mathcal{C}}^2} \right| \right)^{-\frac{1}{3}} \text{Ai}[z] \end{aligned} \right\} \quad (5.12-8)$$

where z is given by

$$\left. \begin{aligned} z &= \Psi'_{v_{\mathcal{C}}} \left(\frac{2}{|\Psi'''_{v_{\mathcal{C}}}|} \right)^{\frac{1}{3}} \text{sign}[\Psi'''_{v_{\mathcal{C}}}] \\ &= \left(\theta_{\rho_{\mathcal{C}}} - \theta_L - 2\tilde{\alpha}_{\mathcal{C}} \right) \left(\frac{1}{2} \left| \frac{d^2 \theta_L}{d\rho_{\mathcal{C}}^2} \right| \right)^{-\frac{1}{3}} \text{sign} \left[\frac{d^2 \theta_L}{d\rho_{\mathcal{C}}^2} \right] \end{aligned} \right\} \quad (5.12-9)$$

In Eq. (5.12-8) the integral form for the Airy function given in Eq. (5.11-8) has been used.

Eq. (5.12-8) tells us a few things. First, near a caustic contact the profile of the signal flaring envelope with orbit angle closely follows the shape of the Airy function of the first kind (first established by George Airy). Figure 5-23 provides a comparison of this third order prediction of amplitude at the LEO

from the nascent (a) and (b) rays with their actual combined amplitude from wave theory for the same case shown in the right side of Figure 5-21 (b). Here the contribution from the main ray (m) is removed. The shape of the third order curve follows the absolute value of the Airy function. The agreement is very close in the neighborhood of the caustic contact point. Third order theory predicts zero amplitude at the nodes, which actually occurs only when the (a) and (b) rays have identical amplitudes and opposite phase. In Figure 5-23 we see slight differences from the wave theory and also a gradual phase misalignment building up as θ_L recedes from the caustic contact. The accuracy of third order theory with decreasing θ_L depends on the degree of symmetry in the θ_L versus ρ_* curve for these two nascent rays about the caustic point. If $|d^3\theta_L/d\rho_*^3|_{\mathcal{C}}$ and the higher derivatives are relatively small, then the defocusing for the two nascent rays will remain roughly equal in magnitude as the impact parameters of the nascent rays move away from the caustic contact point. When close symmetry holds, these rays will continue with decreasing θ_L to mostly null each other at the nodes of the secondary lobes and they will double each other's amplitude at the peaks. The Airy function approximation will be valid in this case over a wider range of θ_L values. Second order geometric optics predicts an infinite amplitude at the caustic contact point at $\theta_L - \theta_o = 4.77$, which is 0.13 mrad greater than the actual location of the peak. But the accuracy of geometric optics improves rapidly away from the caustic contact point. Even at the first node it accurately predicts its amplitude and location.

The second conclusion from Eq. (5.12-8) is that the amplitude of the nascent field at the LEO is proportional $\left(|d^2\theta_L/d\rho_*^2|_{\mathcal{C}}\right)^{-1/3}$, a quantity that is readily obtained from Bouguer's law and the bending angle equation if the refractivity profile is given. We see that caustic flaring comes in two flavors. When $|d^2\theta_L/d\rho_*^2|_{\mathcal{C}}$ is small enough, we will have strong flaring near a caustic contact, on which high frequency fringes from interference with the main ray will cause a minor modulation. Here there is a wider range of spectral numbers associated with the nascent rays that coherently contribute to the spectral integral. An example of this is shown in Section 5.13, Figures 5-31 and 5-32. Rainbow caustics have a similar form. However, when $|d^2\theta_L/d\rho_*^2|_{\mathcal{C}}$ is larger, as is the case in Figure 5-20, then there is a narrower range of spectral numbers contributing coherently to the spectral integral. In this case the amplitude of the nascent field will be smaller; interference fringes from the main ray will become the dominant feature, with the nascent field providing the envelope of the amplitude fringes. Two examples of relatively weak caustics are shown in Figure 5-21.

The third conclusion concerns the location of maximum amplitude. Neither $\Psi''_{\mathcal{C}}$ nor $\Psi'''_{\mathcal{C}}$ is variable; their values are set at the caustic contact point. But, z is a variable and it depends on the value of θ_L . We now adjust θ_L to achieve the maximum value for $|E_{\mathcal{C}}(r_L, \theta_L)|$ given in Eq. (5.12-8). The value $z = 0$ corresponds to the stationary phase point ($\theta_{\rho_{\mathcal{C}}} - \theta_{L\mathcal{C}} - 2\tilde{\alpha}_{\mathcal{C}} = 0$ in the second order stationary phase formulation) also occurring at the caustic point where $1 - 2D_{\mathcal{C}}\tilde{\alpha}_{\mathcal{C}} = 0$, which is the prediction from geometric optics concerning maximum amplitude (i.e., infinity). This condition yields a determination of $\theta_{L\mathcal{C}}$, which is at $\theta_o + 4.77$ for the upper caustic contact point, and at $\theta_o - 12.18$ for the lower point. But, we see that the Airy function achieves greater values away from the stationary phase point at the caustic, which has a value of $\text{Ai}[0] = 0.355$. The maximum value of $\text{Ai}[z]$ nearest to $z = 0$ is 0.536, and it occurs at $z = -1.019$. If we let $\hat{\theta}_L$ be the value of θ_L where $z = -1.019$, then the offset between the geometric optics prediction of the location $\theta_{L\mathcal{C}}$ of maximum amplitude and the third order stationary phase prediction, $\hat{\theta}_L$, is given by

$$\hat{\theta}_L - \theta_{L\mathcal{C}} = 1.019 \left(\frac{1}{2} \left| \frac{d^2\theta_L}{d\rho_*^2} \right| \right)^{-\frac{1}{3}} \text{sign} \left[\frac{d^2\theta_L}{d\rho_*^2} \right]_{\mathcal{C}} \quad (5.12-10)$$

This expression predicts offsets in Figure 5-21 for the upper caustic point at $\theta_L - \theta_o = 4.77$ of $\hat{\theta}_L - \theta_{L\mathcal{C}} = -0.13$ mrad, or $\hat{\theta}_L - \theta_o = -4.64$. At the lower caustic point at $\theta_L - \theta_o = -12.18$, the offset is $+0.16$ mrad, or $\hat{\theta}_L - \theta_o = -12.02$. Figure 5-21 (b) shows that these predictions are very close to the actual offsets obtained from wave theory.

The maximum amplitudes predicted by Eq. (5.12-8) for the nascent field at the LEO is 1.09 near $\theta_L - \theta_o = 4.8$, and 0.90 near $\theta_L - \theta_o = -12$. The actual values are 1.14 and 0.91. Because the amplitudes of the nascent fields are comparable to the main ray, the (m) ray at the upper caustic and the (b) ray at the lower caustic, there will be high frequency fringes of comparable amplitude superimposed on this nascent radiation. Thus, it is the envelope of the field at the caustic points that we should compare. Setting the field amplitude of the main ray to unity, Eq. (5.12-10) predicts the peaks in the envelope to be 2.09 maximum and -0.09 minimum near $\theta_L - \theta_o = 4.7$, and 1.90 and $+0.10$ near $\theta_L - \theta_o = -12$. Since we are displaying amplitudes here the negative minimum peak near $\theta_L - \theta_o = 4.8$ is “reflected” about the θ_L axis and becomes $+0.09$. These compare very closely, if not perfectly, with the actual peaks in the lobe nearest the caustic contact points.

An index to characterize the two flavors of caustic flaring, strong and weak, can be formulated. We take the ratio $E_{\mathcal{C}}(r_L, \theta_L) / E_o \zeta_m^{1/2}$, which is the ratio of the amplitude of the field at the LEO from the nascent (a) and (b) rays at the caustic point $\theta_L = \theta_{L\mathcal{C}}$ divided by the field from the main ray (m) at the same orbit angle. It follows from Eqs. (5.12-6) and (5.12-8) that this ratio is given by

$$\left\{ \mathcal{R} = \frac{E_{\mathcal{C}}}{E_o \zeta_m^{1/2}} = \sqrt{2\pi} \text{Ai}[-1.02] \left| \frac{d\theta_L}{d\rho_*} \right|_{\mathcal{C}}^{\frac{1}{2}} \left| \frac{1}{2} \frac{d^2\theta_L}{d\rho_*^2} \right|_{\mathcal{C}}^{-\frac{1}{3}} \right. \quad (5.12-11)$$

If $|d^2\theta_L / d\rho_*^2|_{\mathcal{C}} \gg 5|d\theta_L / d\rho_*|_{\mathcal{C}}^{3/2}$, then $\mathcal{R} \ll 1$, and vice versa.

Even the fine structure in the minimum envelope in Figure 5-21 (b) at $\theta_L - \theta_o = 4.7$ is predicted well by Eq. (5.12-8). As θ_L is moved away from its maximum amplitude point at $\hat{\theta}_L$, z varies and $\text{Ai}[z]$ diminishes. If we adjust z so that the amplitude of the nascent field at the LEO is unity, this will mark a pair of points along the θ_L axis where the nascent field and the main field from the (m) ray can completely cancel each other. Setting $E_{\mathcal{C}}(r_L, \theta_L) = 1$ yields z values of -1.40 and -0.58 . This yields $\theta_L - \hat{\theta}_L = +0.05, -0.06$, or $\theta_L = 4.69, 4.58$, which corresponds almost exactly with the θ_L values where cancellation is almost complete.

The first node in the envelope nearest the peak corresponds in Eq. (5.12-8) to the first zero of $\text{Ai}[z]$, which occurs at $z = -2.34$. This yields $\theta_L - \hat{\theta}_L = -0.16$, or $\theta_L = \theta_o + 4.48$, which is very close to the location of the actual node in Figure 5-21 (b), $\theta_L = \theta_o + 4.49$. Here third order theory predicts zero for the nodal amplitude, but we see a slight departure from this value. The accuracy of the third order theory deteriorates the further away from the caustic point one applies it. Actually, third order theory does well for several lobes at predicting amplitude peaks and nodes, but eventually it falls out of phase with wave theory results.

The defocusing factors of the (a) and (b) rays at the first node at $\theta_L = \theta_o + 4.49$ are given by $|\zeta_b|^{1/2} = 0.462$ and $\zeta_c^{1/2} = 0.526$. This node marks the point where the (a) and (b) rays first become 180 degrees out of phase. The difference of these defocusing factors is 0.06, which is very close to the actual nodal amplitude at $\theta_L = \theta_o + 4.49$.

One also can predict the location of the first node from geometric optics. We know from Eq. (5.12-2) that at $\theta_L = \theta_{L\mathcal{C}}$ the phase function difference $\Psi_a^* - \Psi_b^* = 0$ because, according to geometric optics, the (a) and (b) rays are merged at that caustic contact point. We now adjust θ_L away from the caustic point until the phase difference is $\Psi_a^* - \Psi_b^* = \pi$, exactly. The expressions for

Ψ^* in Eq. (5.12-2) can be used for this calculation, although care in taking the difference must be exercised because of the enormity of Ψ^* itself. When this program is undertaken for the example shown in Figure 5-21 (b), one obtains $\theta_L = \theta_o + 4.56$ for the location of the node where the (a) and (b) rays have exactly opposite phase, a somewhat less accurate result than the third order stationary phase prediction, but still a rather good prediction.

An alternate approach is to use the third order expansion for Ψ^* around the caustic contact point $(\rho_{\mathcal{C}}, \theta_{L\mathcal{C}})$ in the impact parameter diagram. We use this expansion to calculate the separation distance $\rho_a - \rho_b$ and the change in orbit angle $\theta_{La} - \theta_{Lb}$ required to achieve exactly π radians phase difference between the (a) and (b) rays at the LEO since their nascence. At the caustic contact point $\theta_L = \theta_{L\mathcal{C}}$, and from Eqs. (5.12-5) and (5.12-6) it follows that $\Psi'_{\mathcal{C}} = 0$ and $\Psi''_{\mathcal{C}} = (d\theta_L / d\rho_*)_{\mathcal{C}} = 0$. Therefore, we may write the difference in phase between rays (a) and (b) in terms of the third order expansion as

$$\left. \begin{aligned} \Psi_a - \Psi_b &\doteq \frac{1}{3!} \left(\frac{d^2 \theta_L}{d\rho_*^2} \right)_{\mathcal{C}} \left((\rho_a - \rho_{\mathcal{C}})^3 - (\rho_b - \rho_{\mathcal{C}})^3 \right) \\ &= \frac{1}{24} \left(\frac{d^2 \theta_L}{d\rho_*^2} \right)_{\mathcal{C}} (\rho_a - \rho_b)^3 = \pi \end{aligned} \right\} \quad (5.12-12)$$

It follows that the separation distance in impact parameters of the (a) and (b) rays at their first null since nascence, $\Delta\rho_{*ab}$, is given by

$$\Delta\rho_{*ab} = (\rho_a - \rho_b) \Big|_{\Delta\Psi_{ab}=\pi} = (24\pi)^{1/3} \left| \frac{d^2 \theta_L}{d\rho_*^2} \right|_{\mathcal{C}}^{-1/3} \text{sign} \left[\frac{d^2 \theta_L}{d\rho_*^2} \right]_{\mathcal{C}} \quad (5.12-13)$$

For $\Delta\theta_L = (\theta_L - \theta_{L\mathcal{C}}) \Big|_{\Delta\Psi_{ab}=\pi}$ we obtain

$$\left. \begin{aligned} \Delta\theta_L &\doteq \frac{1}{2} \left(\frac{d^2 \theta_L}{d\rho_*^2} \right)_{\mathcal{C}} (\rho_a - \rho_{\mathcal{C}})^2 \doteq \frac{1}{2} \left(\frac{d^2 \theta_L}{d\rho_*^2} \right)_{\mathcal{C}} (\rho_b - \rho_{\mathcal{C}})^2 \doteq \\ \frac{1}{8} \left(\frac{d^2 \theta_L}{d\rho_*^2} \right)_{\mathcal{C}} (\rho_a - \rho_b)^2 &= \frac{1}{2} (3\pi)^{2/3} \left| \frac{d^2 \theta_L}{d\rho_*^2} \right|_{\mathcal{C}}^{1/3} \text{sign} \left[\frac{d^2 \theta_L}{d\rho_*^2} \right]_{\mathcal{C}} \end{aligned} \right\} \quad (5.12-14)$$

We can compare the coefficients in Eqs. (5.12-9) and (5.12-14) for predicting the location of the first node. The coefficient in Eq. (5.12-9) for $\Delta\theta_L = \theta_L - \theta_{L\mathcal{C}}$ is 2.34, which corresponds to the first zero of the Airy function $\text{Ai}[z]$.

Eq. (5.12-9) predicts the location of the node at $\theta_L = \theta_o + 4.48$, which is very close to the actual first node shown in Figure 5-23. Eq. (5.12-14) gives a coefficient of $(3\pi)^{2/3} / 2 = 2.23$ and it predicts the location at $\theta_L = \theta_o + 4.50$.

We recall that the semi-minor axis of the first Fresnel zone in second order geometric optics is the separation distance of the impact parameters for two rays that arrive at the LEO exactly π radians out of phase. Thus, at a point where $\Psi'' \neq 0$ one obtains

$$\Psi_a - \Psi_b \doteq \frac{1}{2} \Psi'' (\rho_a - \rho_b)^2 = \pi \quad (5.12-15)$$

or

$$k\mathcal{F} = (\rho_a - \rho_b)_{\Delta\Psi=\pi} = \left(\frac{2\pi}{|\Psi''|} \right)^{\frac{1}{2}} = \left| \frac{1}{2\pi} \frac{d\theta_L}{d\rho_*} \right|^{-\frac{1}{2}} = \left\{ \begin{array}{l} \left(\frac{2\pi D_{\rho_*}}{|1 - 2D_{\rho_*} \tilde{\alpha}'_*|} \right)^{1/2} \\ = (2\pi D_{\rho_*} |\zeta_*|)^{1/2} \end{array} \right\} \quad (5.12-16)$$

At any rate, \mathcal{F} is infinite at the caustic point. For comparison, $k^{-1}\Delta\rho_{*ab}$ in Eq. (5.12-13) for the example given in Figure 5-21 at the upper caustic point at $\theta_L = \theta_o + 4.47$ has a value of 880 m. In a vacuum \mathcal{F} is about 750 m at the LEO radial distance of $r_L = 1.1r_o$. In Figure 5-26, which shows the impact parameter diagram for a Gaussian refractivity distribution located in the lower troposphere with a dry air defocusing medium included, $k^{-1}\Delta\rho_{*ab} \approx 290$ m. For dry air at sea level \mathcal{F} is about 250 m. The offset $\Delta\theta_L$ between the primary peak amplitude at the upper caustic point and the first node is about 0.29 mrad in Figure 5-21 and about 0.80 mrad in Figure 5-26. (This corresponds to roughly 300 ms and 800 ms of elapsed time, respectively, plenty of time for a phase-locked loop GPS receiver to decide that it has got the right stuff.)

From these and previous discussions, we conclude that even at the first node after the caustic point, second order ray optics does a good job at predicting the field amplitude, but it gives a somewhat less accurate prediction of the phase. This agreement with wave theory results improves as the impact parameters for the nascent rays recede further from the caustic point.

In summary, geometric optics works well for predicting the amplitude of the field in multipath conditions provided that, 1) the separation in altitudes between impact parameters of the competing rays exceeds a certain minimum distance, and 2) caustic contact neighborhoods are avoided. Condition 1) is equivalent to the requirement in spectral theory that the stationary points in spectral number are sufficiently separated. When this is satisfied the spectral

integrals, one for each stationary phase point, do not significantly co-mingle; that is, the spectral neighborhood providing significant contributions to the spectral integral from one ray does not overlap the spectral neighborhood for the other ray. This is equivalent to saying that the two rays corresponding to these two spectral integrals are separated adequately in impact parameter altitude. An accuracy index based on this separation concept can be derived from second and third order stationary phase theory. This has been given in Appendix D. One can derive from that discussion an accuracy-dependent index for minimum separation altitudes of the impact parameters (or spectral numbers in wave theory) for which second order ray theory will be adequate. For a given accuracy it can be shown that this index is proportional to $(d\theta_L / d\rho_*)^{-1/2}$, or to the first Fresnel zone. From the agreements in amplitude fringes and modulations discussed above, we also can infer that the relative phases between rays are handled accurately by geometric optics away from caustic neighborhoods. Getting absolute phase from ray theory to agree closely with the phase from wave theory is somewhat more challenging because of the extreme sensitivity of absolute phase to boundary conditions; this results from the fact that $r_o / \lambda \gg 1$. But, even the phases between the two systems can be aligned by renormalization.

5.12.5 Reduction of Multipath Spectral Width by Defocusing

The frequencies of the high-frequency interference fringes shown in Figures 5-21 and 5-22 have a temporal equivalent in the range 80-180 Hz. These high frequencies result from using a solitary Gaussian refraction model in these figures. These frequencies exceed by an order of magnitude the actual maximum bandwidth (~10-15 Hz) of transient signatures in the excess Doppler observed by a LEO sounding the Earth's atmosphere. This disparity in the fringe frequency bandwidth appearing in Figures 5-21 and 5-22 versus realistic bandwidths is largely due to the defocusing coming from the gradient of the dry air refractivity, which was omitted in these early figures. In a background medium that is defocusing, the impact parameter spread between multipath rays is greatly compressed. We present now a simple expression that accounts for the defocusing, and from which the qualitative aspects of the interference fringes from a Gaussian model with dry air added can be inferred.

From Bouguer's law given in Eq. (5.6-5) we have

$$\left. \begin{aligned} \theta_L - \theta_o &= \sin^{-1}\left(\frac{\rho_*}{\rho_L}\right) - \sin^{-1}\left(\frac{\rho_o}{\rho_L}\right) - 2(\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_o, \rho_o)), \\ &\doteq D_{\rho_o}^{-1}(\rho_* - \rho_o) - 2(\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_o, \rho_o)), \\ D_{\rho_o} &= \rho_L \cos(\theta_o + 2\tilde{\alpha}(\rho_o, \rho_o)) \end{aligned} \right\} \quad (5.12-17)$$

It follows that the shape of the θ_L versus ρ_* curve, such as that shown in Figure 5-20, is essentially determined by the shape of $\tilde{\alpha}(\rho_*, \rho_*)$.

Suppose that $\tilde{\alpha}(\rho_*, \rho_*)$ is composed of two parts. The first part is from a refractivity gradient for a local medium arising from, for example, a localized water vapor layer. The second part is from a smooth background refractivity gradient arising from, for example, dry air. We write $\tilde{\alpha} = \tilde{\alpha}_w + \tilde{\alpha}_A$ ⁴. We expect that the rapid variations and the caustic points in $\theta_L(\rho_*)$ come from rapid variations in $\tilde{\alpha}_w$. On the other hand, $\tilde{\alpha}_A$ is relatively slowly varying, with a monotonic negative gradient. Therefore, over a sufficiently narrow range of impact parameter values, the linear expression $\tilde{\alpha}_A \doteq \tilde{\alpha}_{Ao} + \tilde{\alpha}'_A(\rho_* - \rho_o)$ will suffice. Because of the non-linearity in $\tilde{\alpha}_w$ versus ρ_* , we have to keep the full expression. In Eq. (5.12-17) we substitute $\tilde{\alpha}_w + \tilde{\alpha}_A$ for $\tilde{\alpha}$, and we hold θ_L fixed while varying ρ_* . This gives us an expression for the breadth $\Delta\rho_*$, or altitude difference at any given epoch between impact parameter values associated with any two competing rays arriving at the LEO. Upon expanding Eq. (5.12-17) through first degree in $\Delta\rho_*$, we have

$$\Delta\rho_*|_{\Delta\theta_L=0} = \rho_{*2} - \rho_{*1} = \frac{2D_{\rho_o}(\tilde{\alpha}_{w2} - \tilde{\alpha}_{w1})}{1 - 2D_{\rho_o}\tilde{\alpha}'_A} \quad (5.12-18)$$

Here $\tilde{\alpha}_{w2} - \tilde{\alpha}_{w1}$ is the difference in one-way bending angles resulting from the local intrusive medium at a given orbit angle θ_L between multipath rays #1 and #2. At sea level, the defocusing factor from dry air is $\zeta_A^{-1} = 1 - 2D_{\rho_o}\tilde{\alpha}'_A \approx 10$. It follows that given the change in bending angle $(\tilde{\alpha}_{w2} - \tilde{\alpha}_{w1})$, $\Delta\rho_*$ and therefore, from Eq. (5.12-4) the multipath Doppler spread $\Delta(d\Psi_*/dt)/2\pi$ will be reduced by about a factor of ten relative to the spread that would be realized without the background refractivity.

5.12-6 Combined Water Vapor and Dry Air Refractivity Model

Figure 5-25 shows $dG^\dagger(\nu)/d\nu$ where Cases a) and b) for the refractivity profile have been combined into the form

⁴ Strictly speaking, we should write $\tilde{\alpha}_w = \tilde{\alpha}_{w+A} - \tilde{\alpha}_A$. In other words, expressing $\tilde{\alpha}_w$ and $\tilde{\alpha}_A$ as separate integrals in Eq. (5.6-5) is not strictly legitimate because of their non-linear dependence on refractivity. Two rays with the same impact parameter value, one in a medium of W+A, and one in a medium of W only, follow different paths. However, for a thin atmosphere Eq. (5.12-18) is a fairly accurate.

$$n-1 = N_o \exp\left(\frac{\rho_o - \rho}{H}\right) + N_w \exp\left(-\frac{(\rho - \rho_w)^2}{2H_w^2}\right) \quad (5.12-19)$$

The curves for dG^\dagger / dv are for different altitudes of the center of the Gaussian distribution near ρ_o ; the center is given by ρ_w . Here $N_w = 0.2N_o$ and $H_w = 0.05H$. The values chosen for the dry air component are: $H = 7k$, $N_o = 270 \times 10^{-6}$ and $\rho_o = 6378(1 + N_o)k$. These yield a refractivity profile at sea level that closely matches the dry air profile for a standard atmosphere. This dry air profile yields a two-sided bending angle at sea level of 20.4 mrad. A marine layer with substantial water vapor that is several hundred meters thick could have a refractivity profile similar to the Gaussian one used in this example. The value used for N_w in this example could correspond to a peak water vapor number density of about 1% of the local density of dry air. This composite moist refractivity profile yields a narrow (~ 1 km) transient with a peak $|n'\rho|$ value of 0.8. This bending angle profile also results in two caustics.

Figure 5-26 is an impact parameter diagram showing the relationship between impact parameter ρ_* and the LEO orbit angle θ_L . This curve results from the combined exponential distribution for dry air plus the Gaussian distribution located in the lower troposphere given in Eq. (5.12-19). Here $k^{-1}\rho_w = 1.75$ km and ρ_o is the impact parameter of a ray with a tangency point at sea level. Also, $\theta_{L_o} = \sin^{-1}(\rho_o / \rho_L) - 2\tilde{\alpha}(\rho_o, \rho_o)$, which is the orbit angle of the LEO at the refracted shadow boundary. For a setting occultation, θ_L decreases with time at a rate of roughly 1 mrad/s. This Figure shows the extensive range ($\Delta\theta_L \sim 20$ mrad, or about 1400 km of LEO orbital motion) or duration (~ 20 s) of multipath at the LEO compared to the half-width H_w (350 m) of the layer. Scintillation in amplitude and phase at the LEO first occurs at ray path tangency altitudes that are about 3 km above the altitude of the layer itself or about 10 s earlier.

Whereas Figures 5-19 through 5-23 use a solitary Gaussian refractivity distribution and include no defocusing effects from dry air, Figure 5-26 includes defocusing effects arising from the background refractivity due to dry air, which manifests itself in the much narrower impact parameter separations for the multiple rays. This defocused scenario produces maximum high frequency fringes of around 15 Hz, much closer to a realistic case. Over most of the multipath zone the bandwidth of the interference spectrum is less than 10 Hz for that example. Defocusing from the dry air refractivity profile compresses the bandwidth of the interference spectrum. But defocusing does not compress the duration of the multipath episode. In Appendix E it is shown that the scale for $(\tilde{\alpha}_{w_2} - \tilde{\alpha}_{w_1})$ for any two competing rays within the multipath zone depends nearly linearly on H_w , the $1-\sigma$ width of the Gaussian distribution,

and it depends only weakly on the peak refractivity value N_w . The width of the multipath zone in orbit angle or in elapsed time, for example, the difference in θ_L between the upper and lower caustic points, depends on $H_w^{-3/2}$; it depends linearly on N_w . As we already have noted in Eq. (5.12-18), the fringe frequency scales, fast and slow, are compressed by defocusing from the slowly varying negative refractivity gradient from the background medium, dry air in this case. This compression ratio from dry air near sea level is about ten-to-one. The excess Doppler difference between rays is $\dot{\theta}_L \Delta \rho_* / 2\pi \approx 5 \Delta r_* \text{ Hz}$, where the difference in tangency points between rays, Δr_* , is in km. Therefore, the fast fringe frequency between the main ray (m) and the nascent rays (a) and (b) shown in Figure 5-26 is about 15 Hz at mid-point in the multipath zone. This may be compared to the fast fringe frequency range of 80-180 Hz in Figure 5-20, which does not include any defocusing from an ambient medium. A simple rule-of-thumb to convert an impact parameter diagram from ρ_* versus θ_L into a Doppler spread versus elapse time is as follows: Multiply the ordinate in km by the factor $5\dot{\theta}_L$ to obtain Hz, and divide abscissa in mrad by $\dot{\theta}_L$ to obtain sec. Here $\dot{\theta}_L$ is the rate of change of the subtending angle between the emitting GPS satellite and the LEO in the plane of propagation, typically 0.6–1.0 mrad/s, mostly depending on an obliquity factor between the LEO orbit and propagation planes.

Figure 5-27 shows the signal amplitude $|E(r_L, \theta_L)|$ at the LEO versus orbit angle $\theta_L - \theta_{Lo}$ during entry into the multipath zone for the same refractivity and orbit model used in Figures 5-26. This Figure does not include the eclipsing by the Earth's limb, including knife-edge diffraction effects. This Figure was obtained from a numerical integration of the spectral integral in Eq. (5.9-10). Above $\theta_L - \theta_{Lo} \approx 8.5$ the main ray (m) provides the principal contribution to $E(r_L, \theta_L)$, without interference from the (a) and (b) rays. For the $\theta_L - \theta_{Lo}$ range shown here, the main ray impact parameter is located 2-4 km above sea level. Because of defocusing from the dry air refractivity gradient there, the amplitude of the (m) ray has been reduced to about 0.4 from its vacuum value of unity. The contact with the caustic surface where $d\theta_L / d\rho_* = 0$ occurs at $\theta_L - \theta_{Lo} \approx 8.54$, but the actual maximum flaring from wave theory occurs at $\theta_L - \theta_{Lo} \approx 8.17$. Geometric optics predicts infinite amplitude at the caustic contact point, but the actual value for the field contribution from the nascent rays (a) and (b) is 0.407. We have already noted that these differences between geometric optics and wave theory predictions can be reconciled using third order stationary phase theory. The fast fringe frequency near maximum flaring is about 13 Hz, which is due to the 2.6 km separation between impact parameters of the (a) or (b) rays at their nascence, from the (m) ray. This rate reduces to a minimum of about 10 Hz near $\theta_L - \theta_{Lo} = 4$. The slow modulation of the envelope, initially at roughly 1 Hz, is

due to interference between the (a) and (b) rays. The node at $\theta_L - \theta_{Lo} = 7.8$ marks the first occasion since their nascence where the (a) and (b) rays are π radians out of phase. As the impact parameter separation distance between the (b) ray and the other two increases, the fast frequency gradually increases to a maximum of about 20 Hz near the lower caustic contact at $\theta_L - \theta_{Lo} = -9.4$.

An additional consequence follows from Figures 5-26 and 5-27. Multipath from a relatively narrow refracting feature low down in the atmosphere, given a sufficient gradient, can be felt by the LEO for nominal ray path tangency points far above (this height difference scales roughly as $H_w^{-3/2}$). For a setting occultation, flaring of the observed signal at some otherwise benign point is a harbinger of things to come. We will return to this case later in connection with the effect of an embedded reflecting surface on the electric field at the LEO.

Finally, Figure 5-28 shows in the thin phase screen the impact parameter altitude h_s versus h_{LG} using the same Gaussian plus dry air refractivity profile used in Figure 5-26, except that $N_w = -0.2N_o$, an unlikely scenario in the lower troposphere. We have discussed the thin screen phase model in Section 5.10. This refractivity profile results in four caustics, five separate ray systems, and also a quasi-shadow zone around $h_s = 1.5$ km as a result of local defocusing there. Figure 5-29 shows the amplitude of the field at the LEO that results from the refractivity profile used to produce Figure 5-28. Multiplying the abscissa by 3 gives the altitude h_{LG} in Figure 5-28. The steep gradient in h_s versus h_{LG} in the thin screen, or equivalently, in the θ_L versus ρ_* curve, results in the shadow zone. The five rays don't concurrently interfere with each other in this example, but only as triplets. But, it would be easy enough to adjust the local refractivity gradient so that the lower caustic point at the nascence of the (d) and (e) rays in Figure 5-28, for example, was raised to an altitude in h_{LG} that was higher than the caustic point at the nascence of the (b) and (c) rays. This would create a quintuplet multipath episode.

5.13 Spectral Coefficients in a Spherical Refracting Medium with an Embedded Discontinuity

The case where the refractivity is discontinuous at $r = r_o$ in a large, homogeneous, spherical symmetric medium was discussed in Chapter 3. Mie scattering [1, 2] forms the basis of that discussion. It employed the stationary phase technique for interpretation and to aid the numerical evaluation of the scattering integrals. All of the scattering equations carry over to the case of a scattering sphere in a spherical symmetric stratified medium if we replace the spectral coefficients for an incident planar wave in those equations in Chapter 3, $i^{l-1}(2l+1)/(l(l+1))$, by the spectral coefficients $a_l^-(\rho)$ for an incoming wave. From Eq. (5.5-21), these spectral coefficients are given by

$$a_l^-(\rho) = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho, v]) \quad (5.13-1)$$

These coefficients are applicable to an initially planar wave that has been subsequently refracted and retarded by the overlying spherical stratified medium before striking the scattering surface.

Let us now consider the spectral coefficients for the outgoing wave, $a_l^+(\rho)$. From Eq. (5.8-24) we have

$$a_l^+(\rho) = i^{l-1} \frac{2l+1}{l(l+1)} \exp\left(-i\left(2G[\rho^\dagger(v), v] - G[\rho, v]\right)\right) \quad (5.13-2)$$

Here the actual functional form of $a_l^+(\rho)$ will depend in part on the physical properties assumed for the refracting and perhaps scattering atmosphere, and also in part on the impact parameter(s) associated with the ray(s). Suppose that the scattering surface is located at $r = r_o$. For $\rho^\dagger > \rho_o$ these coefficients in Eq. (5.13-2) are still applicable. What happens when $\rho^\dagger \leq \rho_o$? That depends on the medium below.

5.13.1 A Medium with a Discontinuity in its Refractive Gradient

We consider a spherical shell with $n(\rho)$ variable for $\rho \geq \rho_o$ and with $dn/d\rho \equiv 0$ for $\rho < \rho_o$. Then Eqs. (5.5-18) and (5.5-22) show that both a_l^- and a_l^+ are constant with ρ when $\rho < \rho_o$. They also must be equal there to avoid the Hankel function singularity at the origin. Recall that the definition of the spherical Bessel function of the first kind is $\psi_l = (\xi_l^+ + \xi_l^-)/2$, which is well-behaved at the origin. It follows in this case that $a_l^+(\rho_o) = a_l^-(\rho_o)$. This also follows from the defining integral for $G[\rho, v]$ given in Eq. (5.5-20), or by its Airy function form given in Eq. (5.7-2), which we use here. Thus, $G[\rho^\dagger, v] = G[\rho_o, v]$, $\rho^\dagger \leq \rho_o$. It follows that the spectral coefficients for the incoming wave are given by

$$a_l^-(\rho) = \begin{cases} i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho, v]), & \rho \geq \rho_o \\ i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho_o, v]), & \rho \leq \rho_o \end{cases} \quad (5.13-3a)$$

At the LEO, $G[\rho_L, v] \equiv 0$ and the spectral coefficients for the outgoing wave are given by

$$a_l^+(\rho_L) = \begin{cases} i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho^\dagger, v]), & \rho^\dagger \geq \rho_o \\ i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho_o, v]), & \rho^\dagger \leq \rho_o \end{cases} \quad (5.13-3b)$$

For a given value of $v = l + 1/2$ such that $\rho^\dagger(v) \leq \rho_o$, $-2G[\rho_o, v]$ is the total phase delay incurred by the l th spectral coefficient of an initially plane wave as a result of that wave passing completely through an intervening refracting shell with an inner boundary at $r = r_o$. When the stationary phase technique is applied to the spectral integrals, it yields a stationary phase value for v that corresponds in geometric optics to a ray with an impact parameter $\rho_* \doteq v^* \leq \rho_o$.

The total field at the LEO consists of both the incoming and outgoing radial components, as given by Eqs. (5.13-3a) for $\rho = \rho_L$ and (5.13-3b), respectively. But, we have already seen in Section 5.9 from our discussion of the stationary phase technique that at the LEO for an occultation only the outgoing components contribute to the scattering integrals given in Eq. (5.9-5). Because the LEO is so far away from a turning point, only the outgoing components yield stationary phase neighborhoods in spectral number. So, we can ignore the incoming components at the LEO for an occultation geometry.

We also note the relative ease with which this spectral technique using osculating parameters can deal with a discontinuity in the gradient of the refractivity at $r = r_o$. This case might correspond to a discontinuity in scale height of the atmosphere, for example, at the boundary of a marine layer, or in the lapse rate of the troposphere, for example, at the tropopause. From Eq. (5.5-20) it follows that $G[\rho, v]$ is continuous with ρ , even though n' is discontinuous at $r = r_o$. Also, $dG[\rho^\dagger(v), v]/dv$ is continuous. It follows that using the stationary phase technique to evaluate the field from the integrals in Eq. (5.9-5), the same stationary phase point in spectral number is obtained as that obtained from the ambient medium without the discontinuity in n' . We know that for a given the position (r_L, θ_L) of the LEO, the stationary phase point v^* is near the impact parameter ρ_* . But the second derivative $d^2G[\rho^\dagger, v]/dv^2$ is discontinuous at $v^* = \rho_o$. This means when using the stationary phase technique that it is necessary to break the integral for the field over spectral number into two parts: one part for $v > \rho_o - \hat{y}^\dagger K_{\rho_o}$, and a second part for $v < \rho_o - \hat{y}^\dagger K_{\rho_o}$. When the impact parameter is close to ρ_o , these two parts when summed interfere in phase, resulting in fringes in the amplitude and phase of the field.

We choose the simple model in Eq. (5.13-4) with which to compare wave theory scattering results with Fresnel diffraction from a thin screen model. Here we assume that $n' \equiv 0$ for $r < r_o$. For $r \geq r_o$ we assume that $n = n(\rho)$, which is arbitrary other than satisfying the thin atmosphere conditions. As a specific example, we assume that n follows an exponential law for $r > r_o$ and is a constant for $r < r_o$; thus

$$\left. \begin{aligned} n^+ &= 1 + N_o \exp\left(-\frac{\rho - \rho_o}{H}\right), \quad \rho \geq \rho_o \\ n^- &= 1 + N_o, \quad \rho \leq \rho_o \end{aligned} \right\} \quad (5.13-4)$$

Figure 5-30 shows the $2dG^\pm/dv$ curves from Eq. (5.7-7) and certain δ curves from Eq. (5.12-1) for this particular refraction profile. From Eq. (5.7-7) it follows that dG^\pm/dv is defined by

$$\left. \begin{aligned} \frac{dG^+}{dv} &= -\pi K_{\rho^\dagger} \int_{\rho^\dagger}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho, \quad \rho^\dagger \geq \rho_o \\ \frac{dG^-}{dv} &= -\pi K_{\rho_o} \int_{\rho_o}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho, \quad \rho^\dagger \leq \rho_o \\ \frac{dG^+}{dv} &= \frac{dG[\rho^\dagger(v), v]}{dv}, \quad \frac{dG^-}{dv} = \frac{\partial G[\rho_o, v]}{\partial v}, \quad \hat{y} \doteq K_{\rho}^{-1}(v - \rho) \end{aligned} \right\} \quad (5.13-5)$$

When $v^* - \rho_o < -\sim 2K_{\rho_o}$, we may replace the Airy functions in these integral expressions for $2dG^-/dv$ by their respective asymptotic forms. From Eq. (5.7-8) we obtain

$$\frac{dG^-}{dv} \doteq -v \int_{\rho_o}^{\infty} \frac{d \log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - v^2}} = \tilde{\alpha}(\rho_o, v), \quad \rho^\dagger \leq \rho_o - K_{\rho^\dagger} \hat{y}^\dagger \quad (5.13-6)$$

Using the refractivity profile given in Eq. (5.13-4), an explicit form $\tilde{\alpha}(\rho_o, v)$ is given in Eq. (5.8-3). Here $\tilde{\alpha}(\rho_o, v)$ is the cumulative bending on a ray at $\rho = \rho_o$ with an impact parameter value of $v \leq \rho_o$. It corresponds to $\alpha^-(v)/2$ given in Section 2.3, Eq. (2.3-10). As we discussed in Section 5.7, the small difference between dG^-/dv and $\tilde{\alpha}(\rho_o, v)$ only shows up in the immediate vicinity of a turning point.

For dG^\pm/dv , we may write its integral in the form

$$2 \frac{dG^+}{dv} \doteq -2v \int_v^\infty \frac{d \log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - v^2}} = 2\tilde{\alpha}(v, v), \quad \rho^\dagger > \rho_o \quad (5.13-7)$$

Here $2\tilde{\alpha}(v, v) = \alpha_L^+(v)$ corresponds to $\alpha^+(v)$ in Eq. (2.3-11), the bending angle for a transecting ray through the upper medium with a value v for its impact parameter such that $v \geq \rho_o$. This ray completely transects the atmosphere unperturbed by the discontinuity lying below.

The δ curves in Figure 5-30 are obtained from Eq. (5.12-1) and they depend uniquely on the LEO angular position θ_L as a parameter when the orbital radius r_L is fixed. From Bouguer's law in Eq. (5.6-5) this dependence may be recast in terms of bending angle, except that in this example bending angle is not unique over a certain range of angular positions. For that reason we have chosen $\alpha_L^+(\rho_*)$, the bending angle in the upper regime, as the parameter. It is unique. By expanding Eq. (5.12-1) in powers of $(v - \rho_*)$, and upon noting that $\partial\theta_v / \partial v = D_v^{-1}$, it follows that over a sufficiently narrow range of spectral numbers $\delta \doteq \alpha_L^+(\rho_*) + (v - \rho_*) / D_{\rho_*}$. The point of first contact with the caustic surface is located at the point of tangency of the $2dG^- / dv$ curve with the δ curve for $\alpha_L^+ = 12.3 \text{ mrad}$, or for $\theta_L - \theta_o = -10.5 \text{ mrad}$ for this example. For LEO angular positions from this value down to $\theta_L = \theta_o - 20.4$, there are three stationary phase points in spectral number and therefore, three bending angles. For orbital angles outside of $-20.4 \leq \theta_L - \theta_o \leq -10.5 \text{ mrad}$, i.e., for δ curves above the uppermost curve or below the lowest one in Figure 5-30, there is just one intersection point, or one stationary phase point and only one value for the bending angle. The LEO-observed phase and amplitude in this region between the caustic contact point and the cusp in bending angles is marked by strong multipath interference effects and by diffraction in the transition regions. In a wave theory context, fringes in the observed amplitude and phase result from interference among the spectral components for spectral numbers in the immediate neighborhood of ρ_o when the impact parameter is nearby. In geometric optics these fringes come from interference among multiple rays.

At the point of tangency of the δ curve with the $2dG^- / dv$ curve, the condition $2d^2G^- / dv^2 = \partial\delta / \partial v$ must hold. The geometric optics equivalent is $d\alpha_L^- / dv = \partial\delta / \partial v = D_v^{-1}$, which is equivalent to the condition that the defocusing factor ζ must be infinite. This wave theory tangency condition in Figure 5-30 is equivalent to the geometric optics condition on the impact parameter (for a circular LEO orbit) $d\rho_* / d\theta_L \rightarrow \infty$, which is shown in Figure 5-31. At this tangency point flaring in the observed signal will be evident, although the actual maximum in flaring occurs at a point slightly offset

from the tangency point. This offset has been discussed in Section 5.12 using third order stationary phase theory.

Figure 5-32 shows the amplitude of the field at the LEO in the vicinity of the first contact with the caustic ($\theta_L - \theta_o = -10.5 \text{ mrad}$) using the same conditions used in Figure 5-30. The flaring is quite prominent in this example. Figures 5-32 through 5-34 were obtained from a numerical integration of the spectral integrals in Eq. (5.9-24) aided by the stationary phase technique to isolate the contributing stationary phase neighborhoods in spectral number. At $\theta_L - \theta_o = -10.5 \text{ mrad}$, a pair of powerful rays, (a) and (b) in Figure 5-31, are created. In this example their individual amplitudes are roughly six times the amplitude of the original (m) ray at this orbital position. Therefore, their contributions to the field at the LEO largely overwhelm the contribution from the (m) ray by an order of magnitude. Third order stationary phase theory is needed at the caustic point in Figure 5-32 to accurately predict the amplitude and location of the peak. Second order geometric optics predicts an infinite amplitude at the caustic point, whereas Figure 5-32 shows that wave theory yields a peak amplitude of about 5.3. In Section 5.12 it is shown that the peak amplitude associated with a caustic is proportional to $(d^2\theta_L / d\rho_*^2)^{-1/3}$, evaluated near the caustic contact point where $d\theta_L / d\rho_* = 0$. This curvature term is readily obtained from Bouguer's law in Eq. (5.12-5), and the bending profile, which is given in Eq. (5.6-5) in terms of the refractivity profile. Eq. (5.12-13) provides gives a ratio for the amplitudes of the nascent rays compared with the amplitude of the main ray, expressed in terms of the first and second derivatives of the curve, θ_L versus ρ_* shown in Figure 5-31. For the specific exponential shell model in Eq. (5.13-4), which we are using in this section, Eq. (5.12-13) yields a flaring from the nascent rays alone that is 12.3 times the amplitude of the (m) ray. This yields a maximum amplitude of 5.3 from all three rays combined, very close to that the actual peak amplitude shown in Figure 5-32 near the caustic contact. The amplitude of the main ray (m) is normalized to unity outside the medium, but defocusing at $\theta_L - \theta_o = -10.5 \text{ mrad}$ has reduced it to 0.4.

These new rays (a) and (b) begin to mutually interfere with each other, as is evidenced by the onset of fringes for $\theta_L - \theta_o \leq -10.5 \text{ mrad}$. As discussed in Section 5.12, Eq. (5.12-4), the frequency of these fringes can be obtained by multiplying the difference in impact parameter values in km by $5\dot{\theta}_L$. Already at $\theta_L = \theta_o - 10.6 \text{ mrad}$ for rays (a) and (b), that frequency has grown to about 40 Hz (with $\dot{\theta}_L = -1 \text{ mrad/s}$). Figure 5-31 shows that the impact parameter values for the (a) and (b) rays promptly separate with decreasing θ_L and their defocusing factors begin to steeply decrease from initially infinite values. This rapid separation in impact parameter values leads to a rapidly increasing fringe frequency with decreasing θ_L . This interference initially results principally

from the mutual interference between the (a) and (b) rays with a smaller modulation from the weaker (m) ray. Later, as shown in Figure 5-33, as the (a) ray becomes defocused, the high frequency interference is between the (b) and (m) rays with a weaker modulation of much low frequency from interference between the (m) and (a) rays.

Figure 5-33 shows the amplitude of the field near the end of the interference zone. Here the (m) and (b) rays are well separated and the anomalous ray (a) has become a minor contributor because of its defocusing (see Figure 5-31). The impact parameter differences read from Figure 5-31 at the end of the interference zone predict very high fringe frequencies here. At $\theta_L - \theta_o = -19$ the frequency of the high frequency fringes between the (m) and (b) rays is about 250 cycles/mrad. The low frequency modulation in Figure 5-33 comes from interference between the (m) and (a) rays, which have narrowly different impact parameter values in this figure. Here $d(\Psi_{*a} - \Psi_{*m})/d\theta_L$ is about 5 cycles/mrad at $\theta_L - \theta_o = -17$, about 2 cycles/mrad at $\theta_L - \theta_o = -19$, and zero at $\theta_L - \theta_o = -20.4$, the end of the multipath zone.

As was also discussed in Section 5.12, the amplitude of the slow modulation in Figure 5-33 can be obtained from the defocusing factor $\zeta^{-1} = 1 - D_{\rho_*} d\alpha_L / d\rho_*$ for each ray. The modulation peaks and valleys are accurately predicted from the four combinations $||\zeta_a|^{1/2} \pm |\zeta_b|^{1/2} \pm |\zeta_m|^{1/2}|$. The slow modulation ($\Delta\theta_L \sim 0.2$ mrad per cycle at $\theta_L - \theta_o = -17$) results from interference between the narrowly separated (m) and (a) rays. The (a) ray becomes very defocused with decreasing θ_L . At $\theta_L - \theta_o = -17$ the amplitude contributions from the three rays based on their defocusing factors have the ratios b:m:a $\sim 1:1/3:1/5$. At $\theta_L - \theta_o = -19$, these ratios are $\sim 1:1/3:1/11$. For $\theta_L - \theta_o < -20.4$, they are 1:0:0.

The mean amplitude of the field for the range of $\theta_L - \theta_o$ values shown in Figure 5-33 is about 1.1. This is about a factor of three greater than the amplitude of the (m) ray here because of the lens-like property of the refracting shell given by Eq. (5.13-4).

Figure 5-34 shows a section of the de-trended phase of the complete field $\Delta\Phi_{mba}^*$ at the LEO expressed in cycles for the range $-11.35 \leq \theta_L - \theta_o \leq -10.85$. It may be compared to the amplitude fringes shown in Figure 5-32, which cover most of this region. This Figure shows the bursts of rapid acceleration in phase that correspond to local neighborhoods where substantial destructive interference occurs, principally between the newly created (a) and (b) rays. These points correspond to the troughs in amplitude in Figure 5-32. Phase changes of $1/2$ cycle occur over a change in θ_L of less than $2 \mu\text{rad}$, or less than 2 ms of elapsed time. For a more realistic refractivity

model with multipath impact parameter separations that are 1/4 the size of those in Figure 5-31, the elapsed time for a 1/2 cycle change from these brief bursts would be closer to 5 ms, which is still a significant operational problem for a closed-loop receiver.

Finally, Figure 5-35 provides a much more benign scenario. It shows the amplitude of the field at the LEO from the spectral integral in Eq. (5.8-1b) for the case where the gradient of the refractivity has a discontinuity of the opposite polarity to that used in Figure 5-32 (and also significantly smaller). The index of refraction for this case is given by

$$\begin{cases} n^+ = 1 + N_o \text{Exp}[-(\rho - \rho_o) / H^+], & \rho > \rho_o \\ n^- = 1 + N_o \text{Exp}[-(\rho - \rho_o) / H^-], & \rho \leq \rho_o \end{cases} \quad (5.13-8)$$

Here $\Delta H = H^+ - H^- > 0$. This profile results in a mild shadow zone in the neighborhood $-6 \leq \theta_L - \theta_o \leq -5$ mrad with minor diffraction effects. The bending angle profile for this example is given in Section 2.3 by the Case B scenario, in particular, Eq. (2.3-21). See also Section 2.4. In this case ΔH is positive. When the tangency point of the ray descends through the surface of the discontinuity, the ray experiences an increased bending rate because of the increased gradient in refractivity in the lower layer. The amplitude profile closely follows the square root of the defocusing function except in the immediate vicinity of the shadow zone. One can compare Figure 5-35 with Figure 2-11 from the thin screen/scalar diffraction approach. Figure 2-11 applies to a discontinuity in the lapse rate at the tropopause of $\Delta(dT/dr) = -7$ K/km, whereas Figure 5-35 applies to a discontinuity in scale height of $\Delta H = +1$ km. There is a rough correspondence between these two quantities in terms of their perturbations on bending angle, which is given by $\Delta H / H \leftrightarrow -H \Delta(dT/dr)$. When this becomes an equality, the perturbations on the bending angle are about the same, and the resulting diffraction pattern is about the same. Here $\Delta H = +1$ km is equivalent to a lapse rate discontinuity of 4 to 5 K/km.

5.13.2 A Transparent Sphere Embedded in a Refracting Medium

For a transparent sphere with a discontinuity in $n(\rho)$ at $\rho = \rho_o$ and $n' = 0$ inside, the spectral coefficients evaluated at the scattering sphere for the total scattering, including the scattering effects of multiple internal reflections, are obtained by modifying the discussion in Section 3.5, Eqs. (3.5-11) and (3.5-15b), to account for the phase delay $-G[\rho_o^+, \nu]$ in the incident wave induced by the refractivity gradient in the overlying medium. The term $b_l^{(0)}$ is called the zeroth degree reflection coefficient because it applies to only an

external reflection from the sphere; $b_l^{(j)}$ is the j th degree reflection coefficient; it applies to a wave that has undergone $j-1$ internal reflections within the sphere. Thus, $j=1$ corresponds to a wave entering and exiting the sphere without undergoing any internal reflections, $j=2$ corresponds to a wave with one internal reflection, and so on. Summing these coefficients over degree yields the total or aggregate scattering coefficient S_l , which is given by

$$S_l = \sum_{j=0}^{\infty} b_l^{(j)} = -a_l^-(\rho_o^+) \left(\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+} \right) \quad (5.13-9)$$

where $\tilde{\mathcal{W}}_l$ is the complex conjugate of \mathcal{W}_l . Here

$$a_l^-(\rho_o^+) = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho_o^+, \nu]), \quad \rho_o^+ = n_o^+ u_o \quad (5.13-10)$$

and

$$\left. \begin{aligned} \mathcal{W}_l &= \xi_l^+(n_o^+ u_o) \xi_l'^-(n_o^- u_o) - (n_o^- - n_o^+) \xi_l'^+(n_o^+ u_o) \xi_l^-(n_o^- u_o), \\ \mathcal{W}_l^\pm &= \xi_l^\pm(n_o^+ u_o) \xi_l'^\pm(n_o^- u_o) - (n_o^- - n_o^+) \xi_l'^\pm(n_o^+ u_o) \xi_l^\pm(n_o^- u_o) \end{aligned} \right\} \quad (5.13-11)$$

Here n_o^+ is the index of refraction on the outer side of the boundary at $r = r_o$; n_o^- is the value on the inner side. For a transparent sphere $n(\rho)$ is constant and real for $\rho < \rho_o$. Upon propagating the aggregate scattering coefficients upward through the refracting medium to the LEO, one obtains

$$S_l(\rho_L) = -i^{l-1} \frac{2l+1}{l(l+1)} \left(\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+} \right) \exp(-i2G[\rho_o^+, \nu]) \quad (5.13-12)$$

Once again, we see that the scattering coefficients are phase delayed relative to their vacuum counterparts by an amount $-2G[\rho_o^+, \nu]$. The total outgoing field is obtained by adding these scattering coefficients given by Eqs. (5.13-10) and (5.13-12) to the spectral coefficients in Eq. (5.13-3) for the direct field. This combination is then used in the spectral integrals in Eq. (5.8-1b).

Rainbow caustics through a refracting atmosphere with a sharp transparent boundary underneath can be obtained from this approach.

5.14 The Scattered Field from a Perfectly Reflecting Sphere Embedded in a Refracting Medium

Continuing the subject of embedded discontinuities begun in Section 5.13, we now consider a perfectly reflecting sphere embedded in a refracting medium. This example has some application to grazing, near-specular reflections from the ocean [9, 10], especially at very shallow angles of incidence where the reflected signal is essentially linearly polarized. The transmitted GPS signals are right-hand circular polarized (RHCP). Therefore, for grazing angles roughly half the original power is potentially available from reflected signals, especially if the sea surface is smooth.

The spectral coefficients $b_l^{(0)}$ for a very large perfectly reflecting sphere in a homogeneous medium are given in Chapter 3, Eq. (3.17-1). Those coefficients were obtained from the more general case of a transparent sphere with a finite discontinuity in the refractivity at its boundary $\rho = \rho_o$. By letting $n \rightarrow \infty$ at $\rho = \rho_o^-$, one obtains in the limit the reflection coefficients. Therefore, in the case of a reflecting sphere in a non-homogenous medium, $G[\rho, \nu]$ is defined only for $\rho \geq \rho_o$. There is essentially infinite phase wind-up in the spectral coefficients traveling inward across that boundary. If we modify the scattering equation in Eq. (3.17-1) to account for the sphere being embedded in a stratified refracting medium we have for the reflection coefficients at the boundary

$$b_l^{(0)}(\rho_o) = -\frac{1}{2} \left(a_l^+(\rho_o) + a_l^-(\rho_o) \frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)} \right), \quad \rho_o = kn_o r_o \quad (5.14-1)$$

At the boundary the spectral coefficients are

$$a_l^+(\rho_o) = a_l^-(\rho_o) = i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho_o, \nu]) \quad (5.14-2)$$

These apply to an initially planar incoming wave. The modification to this equation to account for an initially spherical incoming wave is found in Section 5.3, Eq. (5.5-3ab). We recall from Chapter 3 that the superscript “(0)” on $b_l^{(0)}$ denotes the zeroth degree reflection coefficient, i.e., the coefficient for the ray with only a surface reflection, i.e., no internal reflections within the sphere. This is the only non-zero term for a perfectly reflecting sphere. To obtain $b_l^{(0)}(\rho)$ for $\rho > \rho_o$, we use arguments that are similar to those used in Section 5.8 to propagate the outgoing spectral coefficients through the medium. We define $b_l^{(0)}(\rho_L)$ as the reflection coefficient at the LEO. Assuming that the LEO is out of the medium so that $G[\rho_L, \nu] \equiv 0$, we obtain

$$b_l^{(0)}(\rho_L) = b_l^{(0)}(\rho_o) \exp(-iG[\rho_o, v]) \quad (5.14-3)$$

It follows from Eqs. (5.8-24) and (5.13-1) that the reflection coefficients at the LEO for a perfect spherical reflector in a stratified medium are given by

$$b_l^{(0)}(\rho_L) = -\frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)} \left(I + \frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)} \right) \exp(-i2G[\rho_o, v]) \quad (5.14-4)$$

In other words, for the stratified medium the reflection coefficient of spectral number l at the LEO is delayed in phase by an amount $-2G[\rho_o, v]$ compared to the pure vacuum case discussed in Chapter 3.

The total field at the LEO is obtained from the spectral integrals in Eq. (5.8-1b), which are comprised of the reflection coefficients in Eq. (5.14-4) plus the spectral coefficients for the incident field. For a collimated approaching wave, the spectral coefficients for the incident field are given by

$$\left. \begin{aligned} a_l^-(\rho_L) &= i^{l-1} \frac{2l+1}{l(l+1)} \exp(-iG[\rho_L, v]), \\ a_l^+(\rho_L) &= i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho^\dagger, v]), \quad \rho^\dagger \geq \rho_o \\ a_l^+(\rho_L) &= i^{l-1} \frac{2l+1}{l(l+1)} \exp(-i2G[\rho_o, v]), \quad \rho^\dagger \leq \rho_o \end{aligned} \right\} \quad (5.14-5)$$

Referring to Figure 5-36, we see that for LEO orbital positions above the shadow boundary there are two paths, a direct path with an impact parameter $\rho_d \geq \rho_o = kr_o n(r_o)$, and a reflected path with an impact parameter $\rho_r \leq \rho_o$. For regimes where the stationary phase condition in spectral number, $v^* \doteq \rho_d \leq \rho_o$, would have held in the absence of the reflecting sphere, there are no stationary phase points for v . Here one obtains a diffracted knife-edge decay in amplitude as the GPS satellite becomes occulted by the reflecting sphere.

From Section 5.9 we know that at the LEO the incoming coefficients $a_l^-(\rho_L)$ for the direct ray (d) may be ignored because they provide no stationary phase points in spectral number. The sum of the reflection and outgoing direct coefficients gives the total field at the LEO. They may be recombined into the form

$$\left. \begin{aligned} & b_l^{(0)}(\rho_L) + \frac{1}{2} a_l^+(\rho_L) = \\ & i^{l+1} \frac{1}{2} \frac{2l+1}{l(l+1)} \left(\left(1 + \frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)} \right) e^{-i2G[\rho_o, v]} - e^{-i2G[\rho^\dagger, v]} \right), \quad \rho^\dagger > \rho_o, \\ & \frac{1}{2} i^{l+1} \frac{2l+1}{l(l+1)} \frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)} e^{-i2G[\rho_o, v]}, \quad \rho^\dagger < \rho_o \end{aligned} \right\} \quad (5.14-6)$$

Consider the regime $\rho^\dagger(v) > \rho_o$, the middle line in Eq. (5.14-6). For $v > \rho_o$, $\xi_l'^-(\rho_o)/\xi_l'^+(\rho_o) \rightarrow -1$ rapidly because of the runoff of $\text{Bi}[\hat{y}]$ for $\hat{y} \doteq K_{\rho_o}^{-1}(v - \rho_o) > 0$. Consequently, $b_l^{(0)}(\rho_o) \rightarrow 0$ for increasing $v > \rho_o$. Already at $\hat{y} = +3$, $|\xi_l'^-(\rho_o)/\xi_l'^+(\rho_o) + 1| = 0.001$. This leaves only the term $i^{l-1}(2l+1)/l(l+1) \exp(-i2G[\rho^\dagger, v])/2$ in this regime where $\rho^\dagger(v) > \rho_o$. This is just the spectral coefficient for a direct ray, but applied to a spectral integral that has a lower cutoff at $\rho^\dagger(v) > \rho_o$. For $\rho_d \gg \rho_o$, there will be a stationary phase point v^* well above the cutoff. Assuming that the overlying medium has a monotonic refractivity gradient, i.e., no multipath, there will be only the one stationary phase point near $v = \rho_d$. It follows that when $\rho_d \gg \rho_o$ this integral yields the field from the direct path, unencumbered by the reflection barrier; the lower cutoff at $\rho^\dagger(v) > \rho_o$ has no effect on the value of the integral. This is the field for a wave that has traveled through a refracting medium without influence from the reflecting surface below.

On the other hand, for a LEO orbital position that is lower than that shown in Figure 5-36, for which the direct ray has an impact parameter value $\rho_d \approx \rho_o$, the stationary phase point v^* will be near ρ_o . Here the lower cutoff in the spectral integral near $v = \rho_o$ does affect the calculation of the field at the LEO. It yields the knife-edge diffraction pattern that the wave from the direct ray will exhibit as the GPS satellite appears to approach the limb of the reflecting sphere and becomes eclipsed. If the LEO is located so that $\rho_d \lesssim \rho_o$, then no stationary phase contributions occur for any value of v ; the LEO is in the refracted shadow. We will show these properties later using stationary phase theory.

The spectral coefficients on the lower line in Eq. (5.14-6) where $\rho^\dagger(v) > \rho_o$ provide the principal contribution to the reflected field at the LEO. Referring to Figure 5-36, if the orbital position of the LEO is such that $\rho_d > \rho_o$, then there is also a reflected ray with an impact parameter value $\rho_r < \rho_o$. There will be a stationary phase point in spectral number, $v = v^*$, near $v = \rho_r$, and the spectral integral over a neighborhood about this point provides the principal

contribution to the reflected field. For spectral numbers $\nu < \rho_o$, $\xi_l'^- / \xi_l'^+ \rightarrow -i \exp(-4i(-\hat{y})^{3/2} / 3)$, which also winds up rapidly with decreasing $\nu < \rho_o$. Therefore, for near-grazing reflections only spectral numbers very near but below ρ_o contribute significantly to the spectral integrals for the reflecting part, but they only do so when the reflected ray exists with $\rho_r \leq \rho_o$. Figure 5-38, which we discuss later, shows the impact parameter diagram for both the direct and reflected rays for a reflecting surface embedded in a refracting medium. It shows the narrow separation between ρ_r and ρ_o for near-grazing conditions.

Applying the spectral coefficients for the outgoing part of the wave given in Eq. (5.14-6) to the spectral integrals in Eq. (5.8-1b), one obtain

$$\left. \begin{aligned} E_r^+(\rho_L, \theta_L) &= \frac{E_o}{\rho_L^2} \int_0^\infty \nu^2 \left(\frac{a_l^+(\rho_L)}{2} + b_l^{(0)}(\rho_L) \right) \xi_l^+(\rho_L) P_l^I(\cos \theta_L) d\nu \\ E_\theta^+(\rho_L, \theta_L) &= \frac{E_o}{\rho_L} \int_0^\infty \left(\frac{a_l^+(\rho_L)}{2} + b_l^{(0)}(\rho_L) \right) \xi_l'^+(\rho_L) \frac{\partial}{\partial \theta} P_l^I(\cos \theta_L) d\nu \end{aligned} \right\} \quad (5.14-7)$$

Here the forms for the spectral coefficients for the direct and reflected waves are given in Eqs. (5.14-5) and (5.14-6). The spectral coefficient for the incoming part for the direct wave at the LEO, $a_l^-(\rho_L)$, is excluded, because the LEO is located well into the outgoing region.

Asymptotic forms can be used everywhere in these integrals except in the reflection coefficients for spectral numbers near $\nu = \rho_o$. However, the Airy functions may be used in place of the spherical Hankel functions, $\xi_l'^- / \xi_l'^+ \doteq (\text{Ai}'[\hat{y}] + i \text{Bi}'[\hat{y}] / (\text{Ai}'[\hat{y}] - i \text{Bi}'[\hat{y}]))$.

We now consider the stationary phase points in spectral number in the phasor form of the integrands in Eq. (5.14-7). For an outgoing wave these phasor forms are given by Eqs. (5.9-6) and (5.9-24) but augmented here by the inclusion of the reflection coefficients. We note that $|\xi_l^- / \xi_l^+| = 1$ for real values of l and ρ . Therefore, we may write this ratio in phasor form

$$\frac{\xi_l^-}{\xi_l^+} = \exp(i2\Omega), \quad \Omega = \tan^{-1} \left(\frac{\text{Bi}'[\hat{y}]}{\text{Ai}'[\hat{y}]} \right) \quad (5.14-8)$$

Let $I(\rho_L, \theta_L)$ denote the part of the field at the LEO for a perfectly reflecting sphere that is due to the ratio $\xi_l^-(\rho_o) / \xi_l^+(\rho_o)$. It follows from Eqs. (5.14-6) – (5.14-8) that Eq. (5.9-24) for this part becomes

$$\left. \begin{aligned}
I_r(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta_L}} \int_0^\infty \left(\frac{\sin^3 \theta_v}{\cos \theta_v} \right)^{\frac{1}{2}} e^{i(2\Omega + \Psi_o(+, -))} dv, \\
I_\theta(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta_L}} \int_0^\infty (\sin \theta_v \cos \theta_v)^{\frac{1}{2}} e^{i(2\Omega + \Psi_o(+, -))} dv, \\
\Psi_o(+, -) &= D_v + v(\theta_v - \theta_L) - 2G[\rho_o, v] - \pi/4, \\
D_v &= \sqrt{\rho_L^2 - v^2}, \quad \theta_v = \sin^{-1} \left(\frac{v}{\rho_L} \right), \\
\Omega &= \tan^{-1} \left(\frac{\text{Bi}'[\hat{y}]}{\text{Ai}'[\hat{y}]} \right), \quad \hat{y} \doteq K_{\rho_o}^{-1}(v - \rho_o)
\end{aligned} \right\} \quad (5.14-9)$$

The total direct and reflected field at the LEO is given by

$$\left. \begin{aligned}
E_r(\rho_L, \theta_L) &= I_r(\rho_L, \theta_L) + J_r(\rho_L, \theta_L) \\
E_\theta(\rho_L, \theta_L) &= I_\theta(\rho_L, \theta_L) + J_\theta(\rho_L, \theta_L)
\end{aligned} \right\} \quad (5.14-10)$$

where from Eq. (5.14-6) we have

$$\left. \begin{aligned}
J_r(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta_L}} \int_{v_o^\dagger}^\infty \left(\frac{\sin^3 \theta_v}{\cos \theta_v} \right)^{\frac{1}{2}} (e^{i\Psi^\dagger(+, -)} - e^{i\Psi_o(+, -)}) dv, \\
J_\theta(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho \sin \theta_L}} \int_{v_o^\dagger}^\infty \left(\frac{\sin 2\theta_v}{2} \right)^{\frac{1}{2}} (e^{i\Psi^\dagger(+, -)} - e^{i\Psi_o(+, -)}) dv, \\
\Psi^\dagger(+, -) &= D_v + v(\theta_v - v) - 2G[\rho^\dagger(v), v] - \pi/4, \\
\Psi_o(+, -) &= D_v + v(\theta_v - v) - 2G[\rho_o, v] - \pi/4, \\
v_o^\dagger &\doteq \rho_o + \hat{y}^\dagger K_{\rho_o}, \quad \rho^\dagger(v) \doteq v - \hat{y}^\dagger K_{\rho_o}
\end{aligned} \right\} \quad (5.14-11)$$

The contribution from the $\Psi_o(+, -)$ term to these J integrals mostly will be negligible when $v > \rho_o$. Moreover, because $2\Omega \rightarrow \pi$ rapidly for $v > \rho_o$, this region in the I integral given in Eq. (5.14-9) will essentially completely cancel this $\Psi_o(+, -)$ term in the J integral. But the $\Psi^\dagger(+, -)$ part of the J integrals in Eq. (5.14-11) accounts for the direct field at the LEO. When $\rho_d \gg \rho_o$, this field is undisturbed by the reflecting surface and the stationary phase point in this integral $v^* \gg \rho_o$; but when $\rho_d \approx \rho_o$, $v^* \approx \rho_o$ and J yields a knife-edge diffraction pattern.

5.14.1 Stationary Phase Analysis

Let us now examine the possible stationary phase neighborhoods for the reflection integrals I_r and I_θ given in Eq. (5.14-9). We seek the zero points for $\partial(2\Omega + \Psi_o)/\partial v$. From Eqs. (5.14-8) and (5.14-9) it follows that the stationary phase condition is given by

$$\left. \begin{aligned} \frac{\partial}{\partial v}(2\Omega + \Psi_o) &\doteq -\frac{2\hat{y}}{\pi M[\hat{y}]K_{\rho_o}} + \theta_v - \theta_L - 2\frac{\partial G[\rho_o, v]}{\partial v}, \\ M[\hat{y}] &= \text{Ai}'^2[\hat{y}] + \text{Bi}'^2[\hat{y}], \quad \hat{y} \doteq K_{\rho_o}^{-1}(v - \rho_o) \end{aligned} \right\} \quad (5.14-12)$$

We assume here that the overlying medium has a monotonic refractivity gradient. Therefore, the direct ray system is unique (no multipath). One can show in this case that for LEO orbital angular positions such for $\theta_L + 2\tilde{\alpha}(\rho_o, \rho_o) \geq \theta_{\rho_o} \doteq \theta_o + N_o \tan \theta_o$, where $\theta_o = \sin^{-1}(r_o / r_L)$, that there is a single negative stationary phase point, i.e., $\hat{y}^* < 0$. Otherwise, that point lies in positive territory and the stationary phase contribution from there will largely cancel the $\Psi_o(+, -)$ contribution in the J integral. The latter situation corresponds to the LEO lying in the refracted shadow of the reflecting sphere. The LEO orbital position $\theta_{Lo} = \theta_{\rho_o} - 2\tilde{\alpha}(\rho_o, \rho_o)$ marks the shadow boundary, the beginning (for the setting case) of the eclipse of the GPS satellite by the reflecting sphere. Thus, except in the immediate vicinity of this shadow boundary, the stationary phase value for the spectral number will be less than ρ_o , that is, $v^* \doteq \rho_o + K_{\rho_o} \hat{y}^*$ with $\hat{y}^* < 0$, when $\theta_L + 2\tilde{\alpha}(\rho_o, \rho_o) \geq \theta_{\rho_o}$.

For the J integral essentially only the $\Psi^\dagger(+, -)$ term contributes to the field when combined with the I integral. Except in the interval $v^\dagger \leq v \leq v^\dagger + 3K_{\rho_o}$, the $\Psi_o(+, -)$ term in the J integral rapidly winds up, contributing negligibly to the integral. It follows that to calculate the total field at the LEO using the stationary phase technique, one needs only the contribution to the I integrals in Eq. (5.14-9) from the stationary phase neighborhood provided by the condition in Eq. (5.14-12), plus the $\exp(i\Psi^\dagger(+, -))$ contribution from the J integral in Eq. (5.14-11). For $\rho_d \gg \rho_o$, the latter integral gives essentially the direct field at the LEO unperturbed by the reflecting sphere but refracted by the overlying medium. The spectral treatment for this direct ray has been previously discussed in Section 5.9 and its stationary phase solution is presented there. Note the lower limit ρ_o in spectral number for the integration in Eq. (5.14-11). If the LEO orbital position is such that the impact parameter for a ray unperturbed by the reflecting surface is less than ρ_o , i.e., $\rho_d < \rho_o$, then this J integral will not contribute significantly to the

field. Moreover, if $\theta_L < \theta_{\rho_o} - 2\tilde{\alpha}(\rho_o, \rho_o)$, the I integrals also have no stationary phase points for negative values of \hat{y} , and the contribution from the positive stationary phase points in \hat{y} will be essentially completely canceled by the $\exp(i\Psi_o(+, -))$ term in the J integral. For orbital positions well into this region it will be dark.

If \hat{y} is sufficiently negative, i.e., $\hat{y} \ll -2$, we may use the negative argument asymptotic forms in $M[\hat{y}]$. In this case, $\pi M[\hat{y}] \rightarrow (-\hat{y})^{1/2} (1 - 7/(32\hat{y}^3) + \dots)$. It follows that $2\Omega + \Psi_o(+, -)$ in the I integrals is given by

$$\left. \begin{aligned} 2\Omega + \Psi_o(+, -) &= -\frac{4}{3}(-\hat{y})^{3/2} - \frac{\pi}{2} + \Psi_o(+, -) \\ \text{or} \\ 2\Omega + \Psi_o(+, -) &= 2\left(v\left(\frac{\pi}{2} - \theta_v^o\right) - \sqrt{\rho_o^2 - v^2}\right) + \\ &\quad \sqrt{\rho_L^2 - v^2} + v(\theta_v - \theta_L) - 2G[\rho_o, v] - \frac{3\pi}{4}, \\ \theta_v &= \sin^{-1}\left(\frac{v}{\rho_L}\right), \quad \theta_v^o = \sin^{-1}\left(\frac{v}{\rho_o}\right), \\ \frac{2}{3}(-\hat{y})^{3/2} &= \sqrt{\rho_o^2 - v^2} - v \cos^{-1}\left(\frac{v}{\rho_o}\right), \quad v \leq \rho_o \end{aligned} \right\} \quad (5.14-13)$$

Although the power series expression for \hat{y} in terms of spectral number v and ρ_o for the reflecting sphere, which is given in Eq. (5.14-12), is adequate, its exact form is given in Eq. (5.14-13) for $\hat{y} < 0$. See Eq. (5.4-3) for further discussion. The exact form is useful in a following discussion showing the correspondence between stationary phase in spectral number and the law of reflection.

It follows upon setting $\partial G[\rho_o, v]/\partial v = \tilde{\alpha}(\rho_o, v)$ and setting $\partial(2\Omega + \Psi_o)/\partial v = 0$, that the stationary phase point for $\hat{y}^* < 0$ is given by the condition

$$\left. \begin{aligned} \text{a) } 2\theta_{v^*}^o - \theta_{v^*} + \theta_L + 2\tilde{\alpha}(\rho_o, v^*) &= \pi \\ \text{or} \\ \text{b) } \hat{y}^* &\doteq -\frac{1}{4}K_{\rho_o}^2 \left(\theta_L + 2\tilde{\alpha}(\rho_o, v^*) - \theta_{v^*}\right)^2 \end{aligned} \right\} \quad (5.14-14)$$

To obtain an explicit value for v^* in terms of θ_L , we note that for near-grazing reflections $\theta_L + 2\tilde{\alpha}(\rho_o, v^*)$ will be close in value to θ_o , as shown in Figure 5-37. We expand θ_{v^*} in a Taylor series expansion about θ_{ρ_o} . Here θ_{ρ_o} and θ_L^A are defined by

$$\left. \begin{aligned} \theta_{\rho_o} &= \sin^{-1}\left(\frac{\rho_o}{\rho_L}\right) \doteq \theta_o + N_o \tan \theta_o, \\ \theta_o &= \sin^{-1}\left(\frac{r_o}{r_L}\right), \quad \theta_L^A = \theta_L + 2\tilde{\alpha}(\rho_o, v^*) \end{aligned} \right\} \quad (5.14-15)$$

The angle θ_L^A is the apparent orbit angle for the LEO. Because of the incoming and outgoing bending, $\tilde{\alpha}(\rho_o, v^*)$ on each leg, our coordinate frame will have been effectively rotated clockwise by $2\tilde{\alpha}(\rho_o, v^*)$ (see Figure 5-36). Thus, the angles of incidence and reflection will be less by an amount $\tilde{\alpha}(\rho_o, v^*)$ than they would be for the case of a reflecting sphere in a homogeneous medium.

We rewrite Eq. (5.14-14a) as, $2\theta_{v^*}^o - (\theta_{v^*}^o - \theta_{\rho_o}) + (\theta_L^A - \theta_{\rho_o}) = \pi$. Expanding θ_{v^*} about θ_{ρ_o} in powers of $v^* - \rho_o$, we obtain

$$\theta_{v^*} = \theta_{\rho_o} + \frac{v^* - \rho_o}{\sqrt{\rho_L^2 - \rho_o^2}} + \dots = \theta_{\rho_o} + (\sin \theta_{v^*}^o - 1) \tan \theta_{\rho_o} + \dots \quad (5.14-16)$$

Noting that for near-grazing reflections $\theta_{v^*}^o \approx \pi/2$, it follows from Eq. (5.14-14) that $2\Omega + \Psi_o(+, -)$ has a stationary value in spectral number when $v = v^*$ or $\hat{y} = \hat{y}^*$, which are given by the conditions

$$\left. \begin{aligned} \theta_{v^*}^o &= \frac{\pi}{2} - \frac{\theta_L^A - \theta_{\rho_o}}{2} - \frac{(\theta_L^A - \theta_{\rho_o})^2}{16} \tan \theta_{\rho_o} + \dots, \\ v^* &= \rho_o \sin \theta_{v^*}^o = \rho_o \left(1 - \frac{(\theta_L^A - \theta_{\rho_o})^2}{8} + \dots \right), \\ \hat{y}^* &= -\frac{K_{\rho_o}^2}{4} (\theta_L^A - \theta_{\rho_o})^2 + \dots \end{aligned} \right\} \quad (5.14-17)$$

Because $\theta_L^A = \theta_L + 2\tilde{\alpha}(\rho_o, v^*)$, an iteration is required in Eq. (5.14-17) to determine v^* , once a specific form for $2\tilde{\alpha}(\rho_o, v)$ is given. For the specific exponential refractivity profile used in the numerical examples here for dry air, $2\tilde{\alpha}(\rho_o, v)$ for the reflected ray is given accurately by Eq. (5.8-3). It can be shown that in this case

$$\left. \begin{aligned} 2\tilde{\alpha}(\rho_o, v^*) &= 2\tilde{\alpha}_o - \frac{\beta}{1+\beta} (\theta_L + 2\tilde{\alpha}_o - \theta_{\rho_o}) \times \\ &\left(1 + \frac{\rho_o}{24H(1+\beta)^3} (\theta_L + 2\tilde{\alpha}_o - \theta_{\rho_o})^2 + \dots \right), \\ \tilde{\alpha}_o &= \tilde{\alpha}(\rho_o, \rho_o), \quad \beta = \frac{N_o \rho_o}{H} \end{aligned} \right\} \quad (5.14-18)$$

This expression is accurate to 1% for near-grazing reflections, i.e., $|\theta_L + 2\tilde{\alpha}_o - \theta_{\rho_o}| < 0.05$. It follows for this case that

$$v^* \doteq \rho_o \left(1 - \frac{1}{8} \left(\frac{\theta_L + 2\tilde{\alpha}_o - \theta_{\rho_o}}{1+\beta} \right)^2 \right) \quad (5.14-19)$$

Figure 5-38 shows the impact parameter diagram for this reflection case. This Figure uses the same refractivity model for the overlying medium as that used in Figures 5-30 through 5-34. A circular LEO orbit was used with $r_L = 1.1r_o$. Therefore, the (d) curve in Figure 5-38 for the direct ray is identical to the (m) curve in Figure 5-31. The impact parameters for the (d) and (r) rays merge at $\theta_L - \theta_{\rho_o} = -20.4$, the grazing point. Divide impact parameter value given in Figure 5-38 for the reflected ray (r) by 100 to obtain the correct value. The reflected ray is very defocused for near-grazing conditions. At $\theta_L - \theta_{\rho_o} = 0$ it is about an order of magnitude more defocused than the (d) ray. This can be calculated by recalling from geometric optics that for a circular LEO orbit the defocusing factor is given by $\zeta = |D_{\rho_*} (d\theta_L / d\rho_*)|^{-1}$. If we set $v^* = \rho_*$ in the stationary phase condition given in Eq. (5.14-14a), then we can form the derivative $d\theta_L / d\rho_*$ while still satisfying the stationary phase condition as θ_L and ρ_* vary. We obtain for the defocusing factor ζ_r for the (r) ray

$$\zeta_r^{-1} = D_{\rho_*} \left| \frac{d\theta_L}{d\rho_*} \right|_r^{-1} = \frac{2 \cos \theta_{\rho_*}}{\sin((\theta_L^A - \theta_{\rho_*})/2) \sin \theta_{\rho_o}} - 1 + 2D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*} \quad (5.14-20a)$$

For near-grazing reflections the term $(\theta_L^A - \theta_{\rho_*})$ is small. Also, $\theta_{\rho_*} - \theta_{\rho_o} \doteq -(\theta_L^A - \theta_{\rho_o})^2 / 8D\rho_o$, which is a very small term. Expanding about $(\theta_L^A - \theta_{\rho_o})$ yields

$$\zeta_r^{-1} \doteq \frac{4 \cot \theta_{\rho_o}}{\theta_L + 2\tilde{\alpha}(\rho_o, \rho_*) - \theta_{\rho_o}} - \left(1 - 2D\rho_o \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*} \right) \quad (5.14-20b)$$

The first term on the RHS is dominant for near-grazing reflections. At $\theta_L - \theta_{\rho_o} = 0$ for the case shown in Figure 5-38, the values of the first and second terms are about 120 and 20, respectively. Thus, Eq. (5.14-20) predicts about 0.08 for the amplitude $E_o \zeta_r^{1/2}$ of the reflected ray at $\theta_L - \theta_{\rho_o} = 0$. The amplitude of the (d) ray there is about 0.6, which decreases to less than 0.4 at the grazing point, $\Delta\theta_L = 20$ mrad below.

Figure 5-39 shows the impact parameter diagram for the reflected and direct rays for $\theta_{\rho_o} - 2\tilde{\alpha}(\rho_o, \rho_o) \leq \theta_L \leq \pi$, which covers the entire range of LEO orbit angles for which a reflected ray exists, from the shadow boundary to a vertical reflection. Here Eq. (5.14-14a) has been used to solve for ρ_r using the same refractivity and orbit models that are used in Figure 5-39a.

We note that $(\theta_L^A - \theta_{\rho_o})K_{\rho_o}$ has a scale-invariance provided that third and higher order terms in $(\theta_L^A - \theta_{\rho_o})$ given for \hat{y}^* in Eq. (5.14-17) can be ignored. The stationary phase points for a reflecting sphere of another radius $\tilde{\rho}_o$ is obtained from our problem merely by applying the scale factor $(\tilde{\rho}_o / \rho_o)^{1/3}$ to the results obtained here for $(\theta_L^A - \theta_{\rho_o})K_{\rho_o}$. One of the practical aspects of this is that the *a priori* value for $\tilde{\rho}_o$, for example, in the case of the topography of the ocean surface, is uncertain to some extent in an actual observation sequence of reflected and direct signals. This is particular true for near-grazing conditions. This scale-invariance property may be useful.

For $\hat{y} < 0$ we can determine from the I integrals in Eq. (5.14-9) using the stationary phase technique the contribution from the stationary phase point to the field at the LEO from the reflecting sphere. We need the second derivative $\partial^2(2\Omega + \Psi_o) / \partial v^2$ evaluated at the stationary phase point given by either Eq. (5.14-14) or by Eq. (5.14-17). This is given by

$$\left(D_v \frac{\partial^2(2\Omega + \Psi)}{\partial v^2} \right)_{v^*} = \left(\frac{-2D_v}{\rho_o \cos \theta_v^o} + 1 - 2D_v \frac{\partial^2 G}{\partial v^2} \right)_{v^*} \quad (5.14-21a)$$

Replacing $(\partial^2 G / \partial v^2)_{v^*}$ with $\partial \tilde{\alpha} / \partial \rho_*$ and also $\theta_{v^*}^o$ with its expansion in terms of $(\theta_L^A - \theta_{\rho_o})$, which is given in Eq. (5.14-27), we obtain

$$\left(\frac{\partial^2 (2\Omega + \Psi)}{D_v \partial v^2} \right)_{v^*} \doteq \frac{-4 \cot \theta_{\rho_*}}{\theta_L + 2\tilde{\alpha}(\rho_o, \rho_*) - \theta_{\rho_*}} + 1 - 2D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*} \quad (5.14-21b)$$

which is the same as the defocusing factor from geometric optics given in Eq. (5.14-20b). For orbital positions near but above the shadow boundary ($\hat{y}^* < 0$), it follows from the stationary phase technique that the field at the LEO from the reflecting sphere is given by

$$\left. \begin{aligned} I &\doteq \frac{E_o \exp(i\Phi^*)}{\sqrt{-2\pi i \rho_L \cos \theta_{\rho_o}}} \int_0^\infty \exp \left(i \frac{1}{2} \frac{\partial^2 (2\Omega + \Psi_o)}{\partial v^2} \Big|_{v^*} (v - v^*)^2 \right) dv \\ &\doteq \frac{E_o \exp(i\Phi^*)}{\sqrt{-2\pi i \rho_L \cos \theta_{\rho_o}}} \left(\frac{1}{\pi} \left| \frac{\partial^2 (2\Omega + \Psi_o)}{\partial v^2} \Big|_{v^*} \right)^{-1/2} \int_{-\infty}^\infty \exp \left(-i \frac{\pi}{2} x^2 \right) dx \end{aligned} \right\} \quad (5.14-22)$$

or

$$I(\rho_L, \theta_L) \doteq E_o \exp(i\Phi^*) \sqrt{\zeta_r^*}$$

where

$$\left. \begin{aligned} (\zeta_r^*)^{-\frac{1}{2}} &= \left(D_v \left| \frac{\partial^2 (2\Omega + \Psi_o)}{\partial v^2} \Big|_{v^*} \right| \right)^{\frac{1}{2}} \\ &\doteq \sqrt{\left| \frac{4 \cot \theta_{\rho_*}}{\theta_L + 2\tilde{\alpha}(\rho_o, \rho_*) - \theta_{\rho_*}} - \left(1 - 2D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*} \right) \right|}, \\ \Phi^* &= (2\Omega + \Psi_o - \pi/4)^* \\ &= -\pi + \sqrt{\rho_L^2 - v^{*2}} - 2\sqrt{\rho_o^2 - v^{*2}} + 2(v^* \tilde{\alpha}(\rho_o, v^*) - G[\rho_o, v^*]) \end{aligned} \right\} \quad (5.14-23)$$

and also

$$\left. \begin{aligned} \theta_{\rho_o} &= \sin^{-1}\left(\frac{\rho_o}{\rho_L}\right) \doteq \theta_o + N_o \tan \theta_o, \quad \theta_o = \sin^{-1}\left(\frac{r_o}{r_L}\right), \\ \theta_L^A &= \theta_L + 2\tilde{\alpha}(\rho_o, v^*), \quad v^* = \rho_o \left(1 - \frac{(\theta_L^A - \theta_{\rho_o})^2}{8} + \dots\right), \\ \rho_o &= kn_o r_o, \quad \rho_L = kr_L, \quad D_{\rho_o} = \sqrt{\rho_L^2 - \rho_o^2} \end{aligned} \right\} \quad (5.14-24)$$

The stationary phase term Φ^* in Eq. (5.14-23) gives the stationary value of the phase at the LEO from the reflecting sphere relative to the phase of the field (neither refracted nor reflected) at the center line at $\theta = \pi/2$ in Figure 5-43 (a). The π term gives the phase reversal from the reflection. The term $\sqrt{\rho_L^2 - v^{*2}} - 2\sqrt{\rho_o^2 - v^{*2}}$ accounts for the difference in optical path length relative to the centerline in Figure 5-43 (a). The first term $\sqrt{\rho_L^2 - v^{*2}}$ gives the length along a straight line from the tangent point on the circle of radius v^* centered at the origin to the LEO. The second term $-2\sqrt{\rho_o^2 - v^{*2}}$ subtracts the length ($\sqrt{\rho_o^2 - v^{*2}}$) between that tangent point and the reflection point along this straight line, and it also subtracts the distance ($\sqrt{\rho_o^2 - v^{*2}}$) that the reflection point has moved counterclockwise from the centerline. The third term $2v^* \tilde{\alpha}(\rho_o, v^*)$ accounts for the extra path length along an arc of radius v^* resulting from the refractive bending. The fourth term $-2G[\rho_o, v^*]$ accounts for the extra roundtrip delay from infinity down to a radius v^* resulting from the refractive gradient in the overlying medium. These third and fourth terms may be rewritten in the form, $2(v^* \tilde{\alpha}(\rho_o, v^*) - G[\rho_o, v^*]) \doteq 2(\rho_o \tilde{\alpha}(\rho_o, v^*) - G[\rho_o, \rho_o])$.

The form of the amplitude signature $\sqrt{\zeta_r}$ given in Eq. (5.14-23) is dominated by the $(\theta_L^A - \theta_{\rho_o})^{-1}$ term for near-grazing reflections. This dominance continues with increasing θ_L until the term $(1 - 2D_{\rho_r} \partial \tilde{\alpha} / \partial v)$ in $\partial^2(2\Omega + \Psi_o) / \partial v^2$ given in Eq. (5.14-21) becomes dominant. Equating terms yields a threshold of $\theta_L^A - \theta_{\rho_o} \approx 4\zeta^* \cot \theta_{\rho_o}$, or when θ_L becomes about a quarter of a radian greater than $\theta_o - 2\tilde{\alpha}(\rho_o, \rho_o)$. Here the impact parameter of the direct ray is well above the atmosphere. For impact parameters above this threshold, $\zeta_r^{-1} \rightarrow 1 - 2D_v \partial \tilde{\alpha}(\rho_o, v) / \partial v$ is the defocusing factor at the reflecting surface for a ray with an impact parameter value of v . In this case the

amplitude of the reflected signal at the LEO reduces to $|I| \rightarrow E_o \zeta_r^{1/2}$, the same form that the amplitude of the direct ray has except that the defocusing factor is based on the bending angle $\tilde{\alpha}(\rho_o, v^*)$ for the reflected ray rather than $\tilde{\alpha}(\rho_*, \rho_*)$ for the direct ray. However, polarization effects from reflections at these steeper angles must be considered. The GPS receiver aboard the LEO is configured for RHCP.

5.14.2 Results from Wave Theory

Figures 5-40 through 5-42 show the amplitude of the field at the LEO from a GPS satellite that is being occulted by a perfectly reflecting sphere embedded in a refracting atmosphere. Figure 5-40 shows the amplitude $E_{(d)}$ from the direct ray only. This corresponds to the (d) ray in the impact parameter diagram shown in Figure 5-38. Figure 5-41 shows the amplitude $E_{(r)}$ from the reflected ray (r) only, and Figure 5-42 shows the amplitude of the complete field. These figures show the transition over about 25 mrad in orbit angle, or roughly 1/2 minute in time, as the LEO moves into the shadow. Here an exponential profile (Eq. (5.8-2a)) has been used for the overlying medium with $N_o = 270 \times 10^{-6}$, $k^{-1}H = 7$ km, $r_o = 6378$ km, and $k = 3.31 \times 10^4$ km $^{-1}$. Also, $r_L = 1.1r_o$ and $\theta_o = \sin^{-1}(r_o / r_L)$. From Eqs. (5.6-5) or (5.8-3) one can show that $2\tilde{\alpha}(\rho_o, \rho_o) = 20.4$ mrad for these parameter values. Figures 5-40 through 5-42 were obtained from a numerical integration of Eq. (5.8-1b) aided by the stationary phase technique. The impact parameter diagram in Figure 5-38 effectively provides those stationary phase points in spectral number for a given value of θ_L .

The mean amplitude in Figure 5-40 shows a rapid decay for θ_L values such that $\rho_d < \rho_o$, or $\theta_L - \theta_{\rho_o} < -2\tilde{\alpha}(\rho_o, \rho_o)$, the penumbra region. The GPS satellite being eclipsed by the reflecting sphere causes the knife-edge diffraction pattern in Figure 5-40. For $\rho_d > \rho_o$, Figure 5-41 clearly shows the $\sqrt{\zeta_r}$ signature in the mean amplitude for the reflected wave that is predicted from the stationary phase technique in Eq. (5.14-23). This agrees with the geometric optics prediction in Eq. (5.14-20) for LEO orbit angles above the shadow boundary, $\theta_L > \theta_{\rho_o} - 2\tilde{\alpha}(\rho_o, \rho_o)$.

Figure 5-42, which gives the wave theory prediction of the total field at the LEO over the same near-grazing orbit angles shown in Figures 5-40 and 5-41, shows the fringes from interference between the direct and reflected rays. Very high fringe frequencies develop as the separation in altitude between the impact parameters of the reflected and direct rays increases with increasing θ_L . The single-sided amplitude of the interference fringes here nearly equals the amplitude of the reflected ray (r) in Figure 5-41. The fringe amplitude would be

significantly reduced for an imperfectly reflecting sphere. The fringe frequency here ranges from about 60 Hz at $\theta_L - \theta_{\rho_o} = 0$, about 20 s before eclipse for the setting case, to zero at $\theta_L - \theta_{\rho_o} = -2\tilde{\alpha}(\rho_o, \rho_o) = -20.4$ mrad, at the refracted shadow boundary.

5.14.3 Law of Reflection

Chapter 3 provides a geometric optics interpretation of the stationary phase point for a reflecting sphere in a homogeneous medium. There it was shown that when the spectral number assumes its stationary phase point, it corresponds to the actual reflection point on the sphere where the law of reflection holds. We also can similarly interpret the stationary phase results for a reflecting sphere embedded in a refracting medium. Referring to Figure 5-43, which is a representation in spectral number space, we have two concentric circles in each panel. The outer circle corresponds to the reflecting sphere and it has a fixed radius ρ_o in spectral number space. The inner one of variable radius is a circle of radius ν , the spectral number. In panel (a), $\nu = \nu^*$, the stationary phase value. In panel (b), the provisional reflection point has been moved clockwise; here $\nu < \nu^*$, and the angle of incidence ζ [QA is the “i” supposed to be in Berthold Script as is the “i” in Eq. (5.14-25)?] is greater than the angle of reflection $\nu = \theta_{\nu}^o$. Noting the triangle ABC in (b), we see that the sum of the interior angles of this triangle satisfies the relationship

$$\zeta + \nu - \theta_{\nu} + \theta_L^A = \pi \quad (5.14-25)$$

for any provisional point of reflection on the circle of radius ρ_o in spectral number. But, at the actual reflection point where $\zeta = \nu$, Eq. (5.14-25) becomes the same relationship given in Eq. (5.14-14a) for the stationary phase condition on the spectral number. We conclude that $\theta_{\nu^*}^o$, which always equals the angle of reflection ν by construction, also equals the angle of incidence ζ when the spectral number assumes its stationary phase value, thus establishing the law of reflection. It follows that

$$\zeta = \nu = \theta_{\nu^*}^o \doteq \frac{\pi}{2} - \frac{\theta_o^A - \theta_{\rho^*}}{2} \quad (5.14-26)$$

Lastly, we see in Figure 5-43 (a) that the position of the reflection point on the reflecting sphere is located at a point counterclockwise relative to the mid-point. This rotation ε , is given by

$$\left. \begin{aligned} \varepsilon &= \frac{\pi}{2} - \theta_{v^*}^o = \frac{\theta_L^A - \theta_{\rho_o}}{2} + \frac{(\theta_L^A - \theta_{\rho_o})^2}{16} \tan \theta_{\rho_o} + \dots \\ &\doteq \frac{\theta_L - \theta_o}{2} + \tilde{\alpha}(\rho_o, v^*) - \frac{1}{2} N_o \tan \theta_o \end{aligned} \right\} \quad (5.14-27)$$

For the exponential refractivity model given in Eq. (5.8-2) and from the resulting bending angle expression given in Eq. (5.8-3), we have

$$\varepsilon \doteq \frac{\theta_L - \theta_o}{2} + \tilde{\alpha}(\rho_o, \rho_o) - \frac{N_o \rho_o}{2H} (\theta_L - \theta_o) - \frac{1}{2} N_o \tan \theta_o \quad (5.14-28)$$

Thus, the refractive bending (for a negative refractivity gradient, i.e., $\tilde{\alpha}(\rho_o, v^*) > 0$) acts to increase the counterclockwise rotation of the reflection point. This is qualitatively indicated in Figure 5-36. For sea level conditions and for $\theta_L - \theta_o = 10$ mrad, the third and fourth terms on the RHS of Eq. (5.14-28) provide a correction of about 10%.

For a backward reflection case, that is, when $\theta_L > \pi/2$, the incoming spectral coefficients $a_l^-(\rho_L)$ also would have to be taken into account. When $\theta_L \gg \pi/2$, this geometry would correspond more closely to a near-vertical reflection geometry rather than the near-grazing one discussed here.

References

- [1] C. Chapman and J. Orcutt, "The computation of body wave synthetic seismograms in laterally homogeneous media," *Rev. Geophys.*, vol. 23(2), pp. 105–163, 1985
- [2] G. Mie, "Beiträge zur Optik Trüber Medien speziell kollidaler Metallösungen," *Ann. d. Physik* (Leipzig), vol. 25, pp. 377–452, 1908
- [3] M. Born and E. Wolf, *Principles of Optics*, 6th ed. Pergamon Press, Oxford, U.K., 1980
- [4] J. Jackson, *Classical Electrodynamics*, 2nd Ed., J. Wiley & Sons, Inc., 1975
- [5] Eckersley 1938 xx
- [6] J. Mathews and R. Walker, *Mathematical methods of Physics*, The Benjamin/Cummings Publishing Co. Menlo Park, CA, 1970
- [7] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, The Institution for Electrical Engrs., London, UK, 2000
- [8] M. Abramowitz, and I. Stegun, ed., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, U.S.

Department of Commerce, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., 1964

- [9] A. Pavelyev, A. Zakharov, A. Kucherjavenkov, E. Molotov, I. Siderenko, I. Kucherjavenkova and D. Pavelyev, "Propagation of radio waves reflected from Earth's surface at grazing angles between a low-orbit space station and a geostationary satellite," *J. Comm., Technology, and Electronics*, vol. 42(1), pp. 51–57, 1997
- [10] G. Beyerle, K. Hocke, J. Wickert, T. Schmidt, C. Marquardt, and Ch. Reigber, "GPS radio occultations with CHAMP: A radio holographic analysis of GPS signal propagation in the troposphere and surface reflections," *J. Geophys. Res.*, vol. 107(D24), 10.1029/2001JD001402, 2002

Figure 5-1. Ray path geometry for a scattering sphere embedded in a stratified medium.

Figure 5-2. Phase rate term $g_l(\rho)$ for spectral numbers near ρ .

Figure 5-3. Solution for refractive bending angle and impact parameter at intersection of curves. Because $2\tilde{\alpha}$ is monotonic the solution is unique.

Figure 5-4. Comparison of the gradient of the wave theory phase function $\partial G[\rho, \nu]/\partial \nu$ with the cumulative refractive bending angle $\tilde{\alpha}$ from geometric optics, in the vicinity of a turning point.

Figure 5-5. Comparison of defocusing quantity $\partial^2 G/\partial \nu^2$ from wave theory with the analogous quantity $\partial \tilde{\alpha}/\partial \nu$ from geometric optics, in the vicinity of a turning point.

Figure 5-6. Variation of $G[\rho, \nu_o]$ with ρ near $\rho = \nu_o$ obtained from eqs. (5.7-1) and (5.7-2). This form is invalid for increasing $\hat{y} > \hat{y}^\dagger$.

Figure 5-7. Gradient of the phase of $a_l^-(\rho)$ in an Airy layer. $\partial \vartheta_l^-/\partial \rho$ is the exact value; $-(d \log n/d\rho)g(\hat{y})$ is from modified Mie scattering theory. Phase units are mrad. (a) $\beta = 0.24$. (b) $\beta = 0.9$.

Figure 5-8. Schematic of simple incoming and outgoing regions.

Figure 5-9. Example of a topology with overlapping incoming and outgoing regions; from a sharp change in refractivity gradient.

Figure 5-10. $\partial G/\partial \nu$ curves for an exponential refractivity profile. Dashed curve shows $\partial G/\partial \nu$ at $\nu = \rho_* + \hat{y}^\dagger K_{\rho_*}$ as ρ_* varies, or equivalently, $\tilde{\alpha}(\rho_*, \rho_*)$.

Figure 5-11. $\partial G/\partial \nu$ curves for a Gaussian refractivity profile for different locations of its center. Dashed curve is the value of $\partial G/\partial \nu$ at $\nu = \rho_* + \hat{y}^\dagger K_{\rho_*}$, or equivalently, $\tilde{\alpha}(\rho_*, \rho_*)$.

Figure 5-12. Variation of $\Gamma(\hat{y})$ in the vicinity of a turning point.

Figure 5-13. Comparison of $U_l(\tilde{\rho})$ from a numerical integration of eq. (5.8-5) in an Airy layer with the osculating parameter solution. (a) Wave equation solutions. (b) Difference between solutions.

Figure 5-14. Comparison of wave equation solutions in a strong refracting medium. (a) Refractivity profile. (b) Bending angle profile. (c) Wave equation solutions, $U_l(\rho)$ and $C_l(a_l^+(\rho)\xi_l^+(\rho) + a_l^-(\rho)\xi_l^-(\rho))$.

Figure 5-15. Comparison of wave equation solutions in a severely refracting medium with a super-refracting layer. (a) Refractivity profile. (b) Bending angle profile. (c) Impact parameter profile. (d) Wave equation solutions; spectral number $\nu = kn(r_o)r_o < kn(r^c)r^c$.

Figure 5-16. Comparison of wave equation solutions with the spectral number located in the super-refracting zone. (a) Wave equation solutions with $\nu = kn(r^c)r^c$. (b) Profile for \hat{y} . (c) Wave equation solutions with $kn(r^c)r^c < \nu < kn(r_u)r_u$. (d) Profile for \hat{y} .

Figure 5-17. Stationary phase solution ν^* to eq. (5.9-13) for the Case a) refractivity profile (eq. (5.8-2a)).

Figure 5-18. Geometric relationships between ray path and phase screens for collimated incident rays.

Figure 5-19. Wave theory multipath zone in spectral number for a Gaussian refractivity profile. The phasor $\exp(i\Psi(+, -))$ in the spectral integral is stationary at an intersection of the δ and $2dG^+ / d\nu$ curves, which gives the bending angle $2\tilde{\alpha}$ of the corresponding ray. Tangent points are caustic contacts where $\partial\Psi / \partial\nu = \partial^2\Psi / \partial\nu^2 = 0$.

Figure 5-20. Impact parameter diagram from geometric optics. This figure shows ρ_* versus LEO orbit angle θ_L in the neighborhood of the same refracting layer described in Figure 5-19. Labels (m), (a) and (b) identify the corresponding rays: main, anomalous, and branching.

Figure 5-21. Amplitude of the field at the LEO versus orbit angle for the same orbit model and Gaussian refractivity profile used in Figures 5-19 and 5-20. (a) Amplitude over the entire multipath zone. (b) Amplitude in the vicinity of the upper and lower caustic contact points.

Figure 5-22. Phase difference near the upper caustic, expressed in cycles of the complete field at the LEO minus the field of the (m) ray.

Figure 5-23. Comparison of the amplitude of the field at the LEO from only the nascent rays (a) and (b) based on wave theory versus third order stationary phase theory. Refractivity and orbit models are the same as those used in Figures 5-19 through 5-22.

Figure 5-24. Quadratic behavior of ρ_* for nascent rays near caustic.

Figure 5-25. Bending angle by a refractivity profile from dry air plus a narrow Gaussian marine layer that is located at different altitudes.

Figure 5-26. Impact parameter diagram with defocusing. LEO orbit radius is $\rho_L = 1.1\rho_o$. This figure shows ρ_* versus θ_L in the multipath zone from a narrow Gaussian refractivity profile from a marine layer, plus a more slowly varying dry air exponential distribution.

Figure 5-27. Amplitude of the field at the LEO versus orbit angle near entry into the multipath zone for the defocused case using the same refractivity and orbit models used in Figure 5-26. The equivalent elapsed time shown in this figure is about 5 s.

Figure 5-28. Impact parameter diagram in the thin phase screen for the same refractivity profile in Figure 5-26, except that $N_w = -0.2N_o$.

Figure 5-29. Amplitude of the field at the LEO for the refractivity and orbit model used in Figure 5-28. Double reversal in the polarity of the bending angle gradient causes four caustics and two multipath zones

Figure 5-30. Stationary phase curves for a spherical shell with a discontinuity in $dn/d\rho$ at r_o per eq. (5.13-4). Stationary phase points occur at the intersections of the δ and $2dG^\pm/dv$ curves.

Figure 5-31. Impact parameter diagram for the refracting shell described in Figure 5-30. Boundary encountered at $\theta_L - \theta_o = -20.4$.

Figure 5-32. Amplitude of the field at the LEO near the caustic contact using same refractivity and orbit models used in Figure 5-30.

Figure 5-33. Continuation of Figure 5-32 to the lower boundary of the interference zone at $\theta_L - \theta_o = -20.4$.

Figure 5-34. De-trended phase at the LEO near onset of multipath. Figures 5-30 through 5-34 use same refractivity and orbit models.

Figure 5-35. Amplitude at LEO from a discontinuity in scale height.

Figure 5-36. Direct and reflected paths for a sphere of radius r_o embedded in a medium with a monotonic refractivity gradient.

Figure 5-37. Geometry for LEO near the shadow boundary of a reflecting sphere embedded in a refracting medium.

Figure 5-38. Impact parameter diagram for a reflecting sphere embedded in a refracting medium versus orbit angle near grazing. The direct curve (d) is identical to the (m) curve in Figure 5-31.

Figure 5-39. Large scale version of the impact parameter diagram over the entire LEO orbit above the grazing point.

Figure 5-40. Perfectly reflecting sphere embedded in a refracting medium. Wave theory prediction of the amplitude $|E_{(d)}(r_L, \theta_L)|$ of the field at the LEO from the direct ray only.

Figure 5-41. Amplitude of the field at the LEO from a perfectly reflecting sphere embedded in a refracting medium; reflected ray (r) only.

Figure 5-42. Amplitude of the complete field (d)+(r) at the LEO from a perfectly reflecting sphere embedded in a refracting medium, the direct ray from Figure 5-40 plus the reflected ray from Figure 5-41.

Figure 5-43. Law of reflection. Geometry in spectral number space for a reflecting sphere embedded in a refracting medium. Outer circle describes reflecting sphere of radius r_o , which maps in spectral number space into a fixed circle of radius ρ_o . Inner circle has a variable radius value, ν . (a) Radius of inner circle is the stationary phase value, ν^ , which gives $\dot{\epsilon} = \dot{\epsilon}$. (b) $\nu < \nu^*$, which gives $\dot{\epsilon} > \dot{\epsilon}$.*

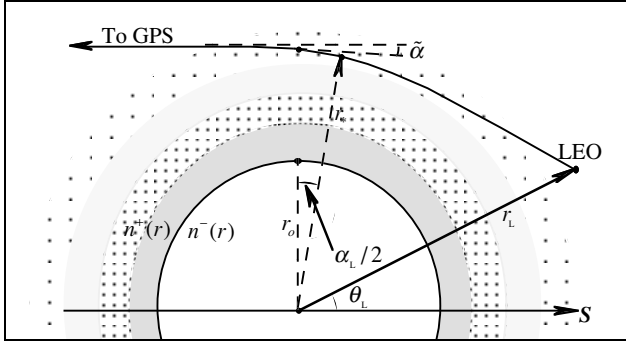


Figure 5-1.

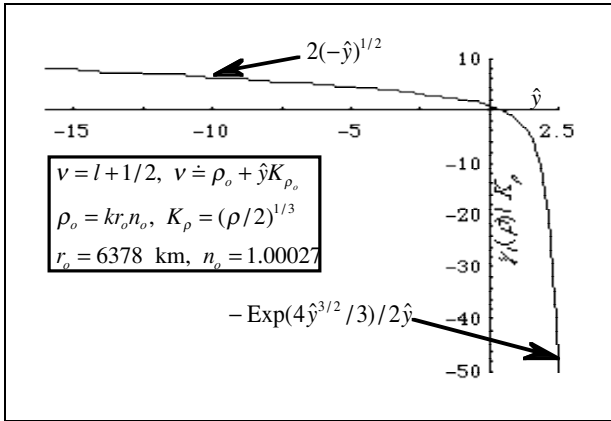


Figure 5-2.

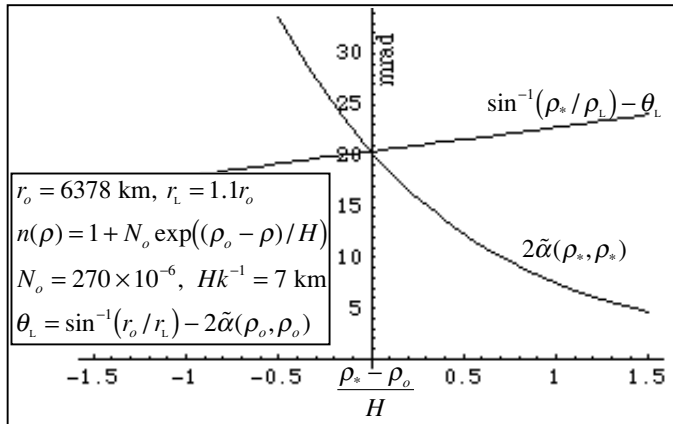


Figure 5-3.

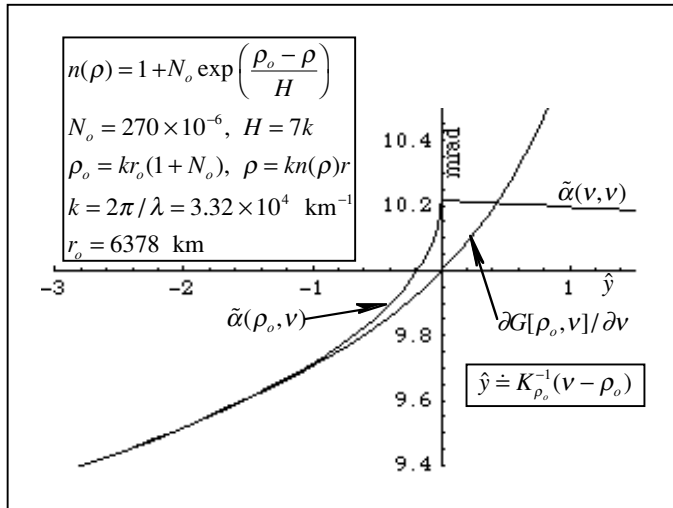


Figure 5-4.

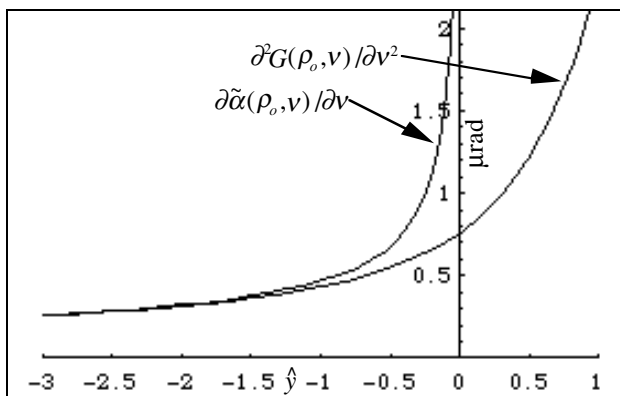


Figure 5-5.

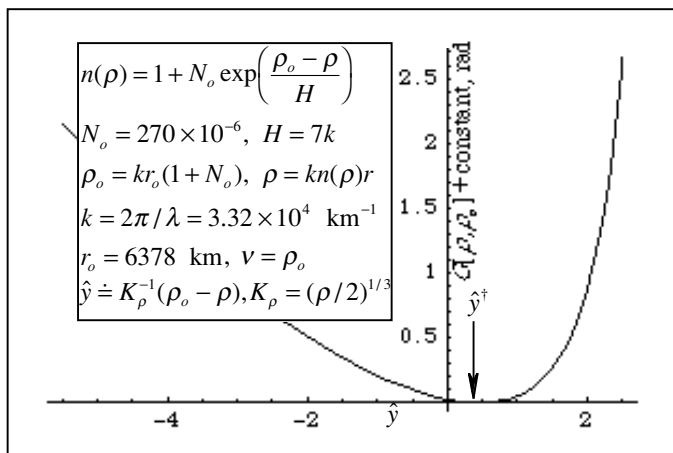


Figure 5-6.

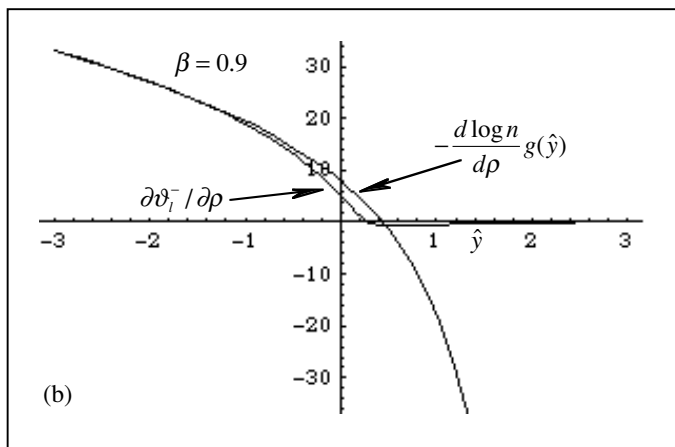
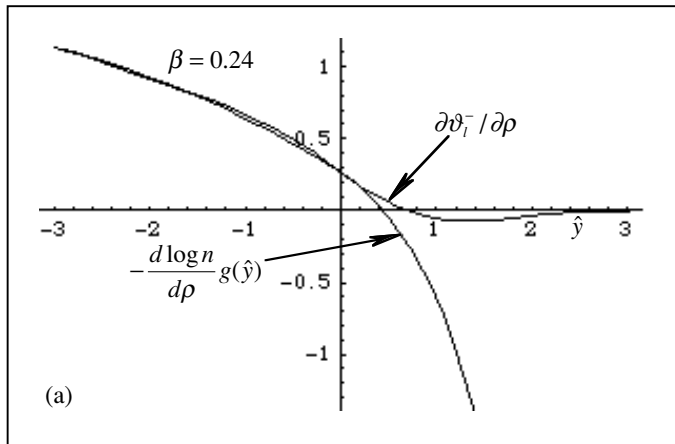


Figure 5-7ab.

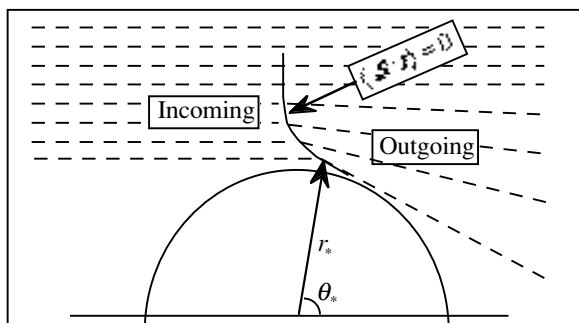


Figure 5-8.

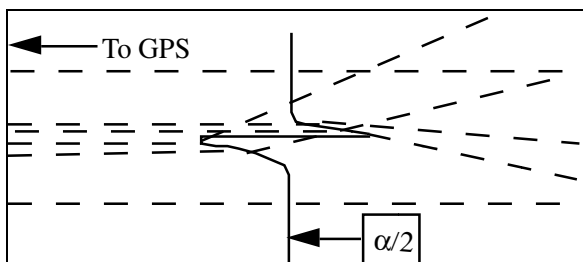


Figure 5-9.

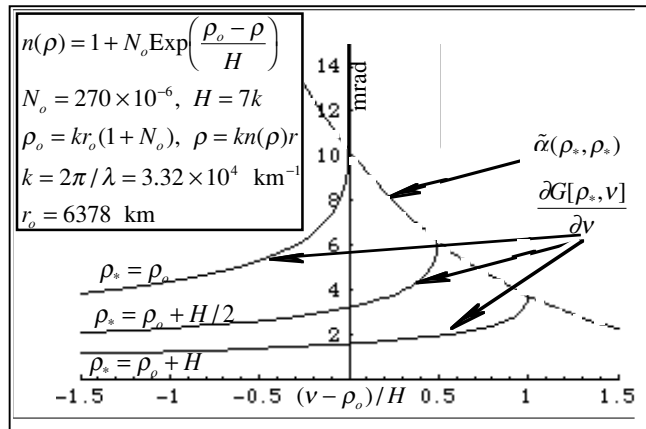


Figure 5-10.

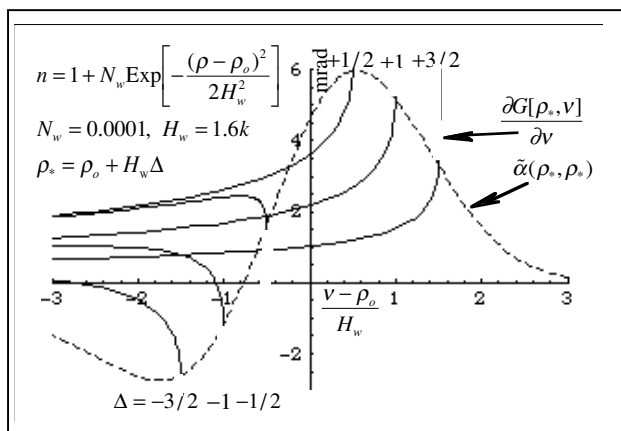


Figure 5-11.

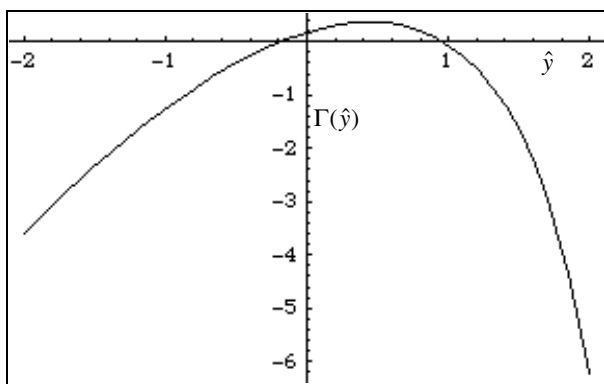
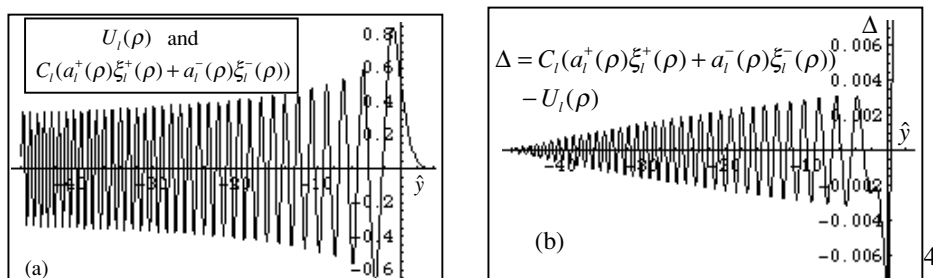


Figure 5-12.

Figure 5-13ab.



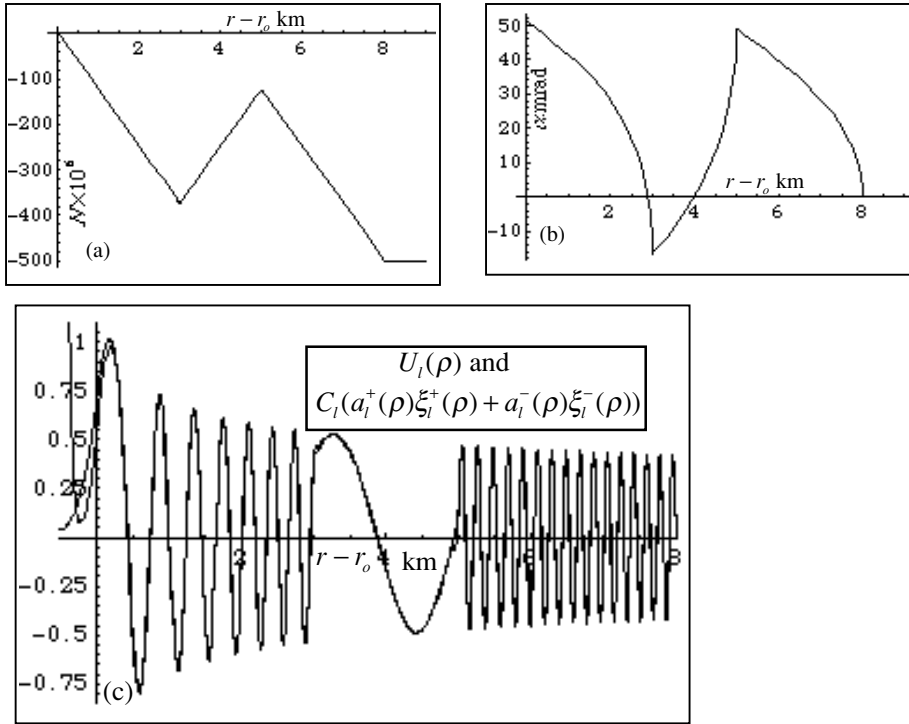


Figure 5-14abc.

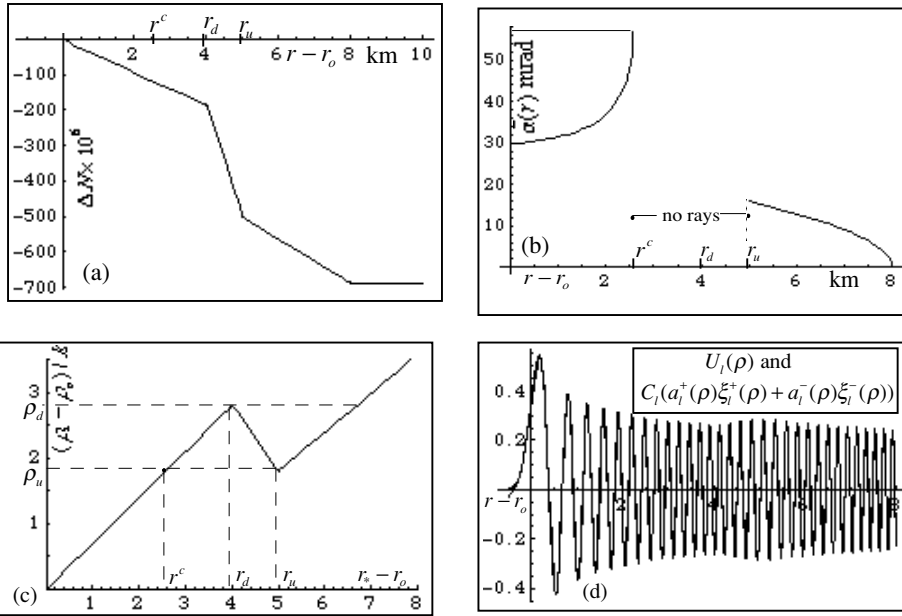


Figure 5-15abcd.

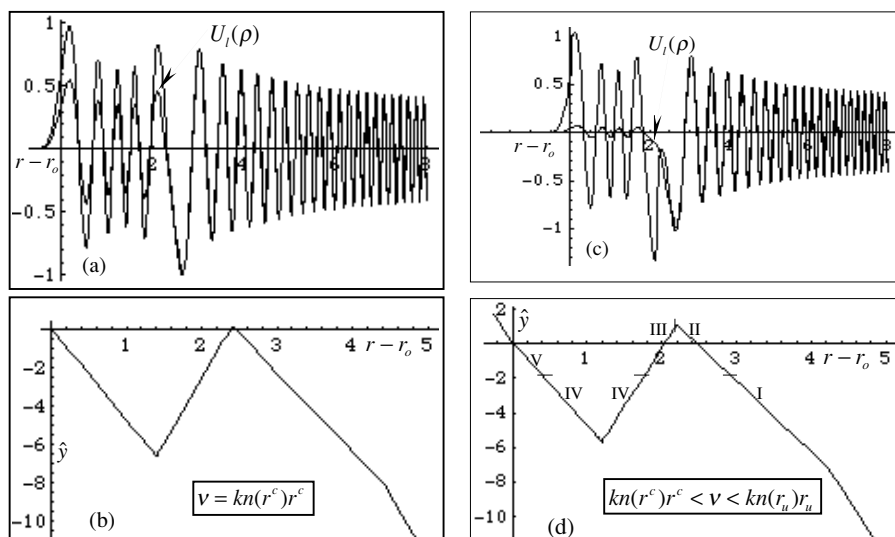


Figure 5-16abcd.

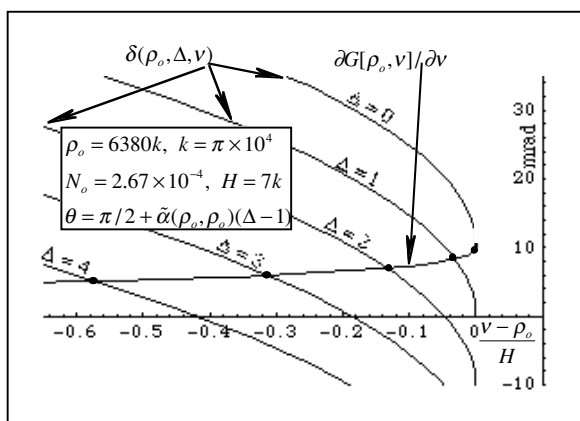


Figure 5-17.

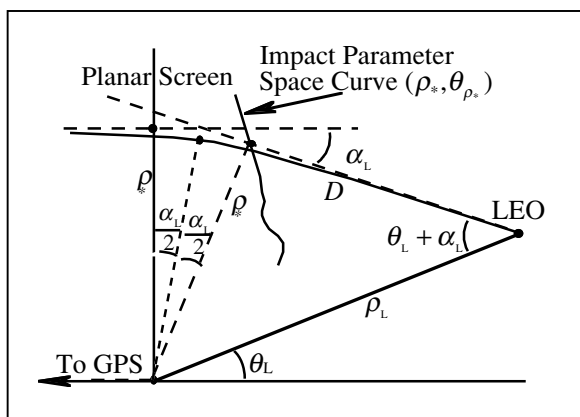


Figure 5-18.

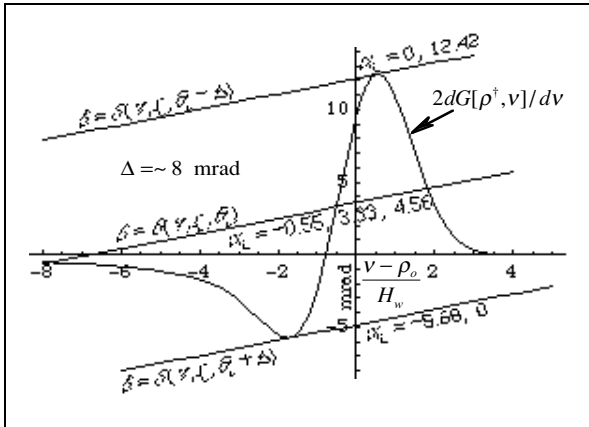


Figure 5-19.

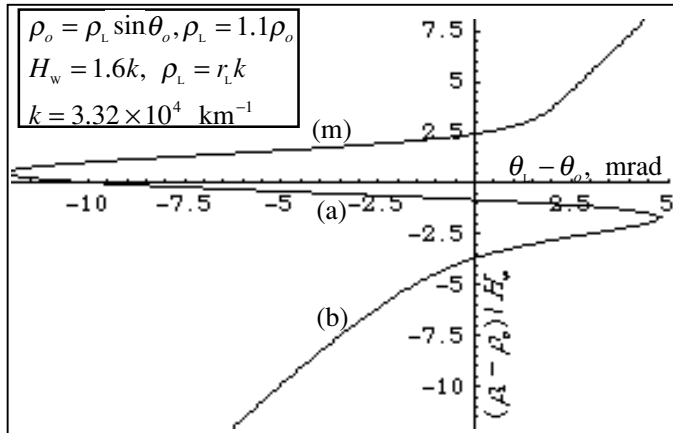


Figure 5-20.

Figure 5-21a.

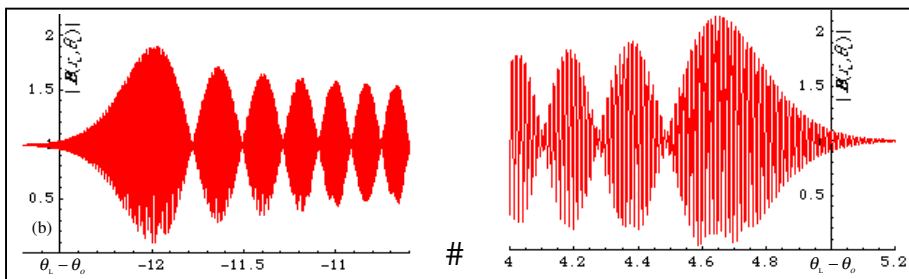
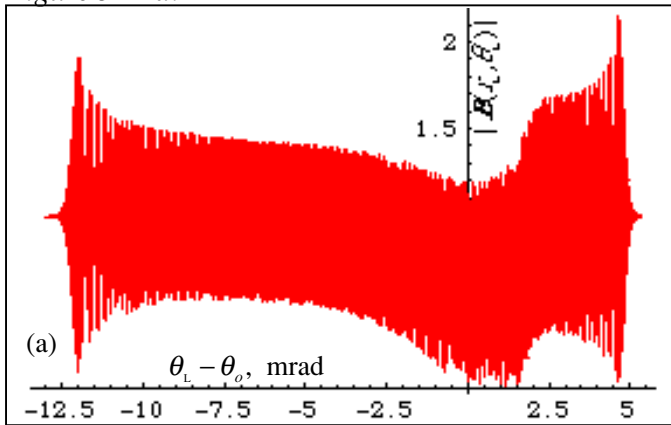


Figure 5-21b.

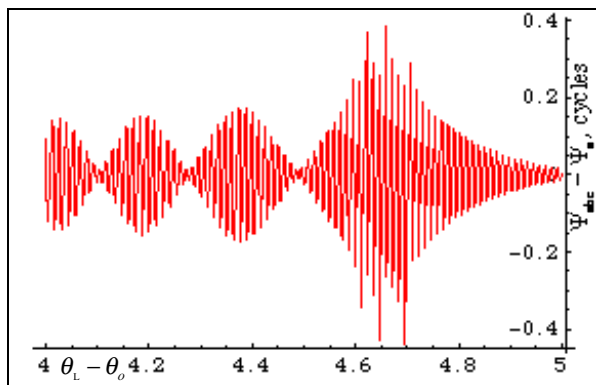


Figure 5-22.

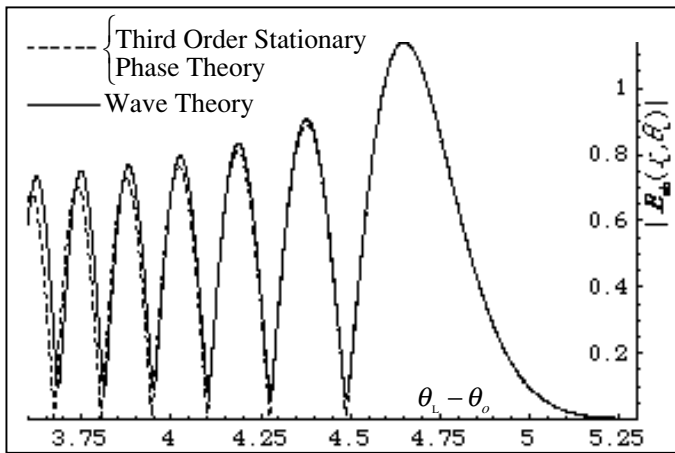


Figure 5-23.

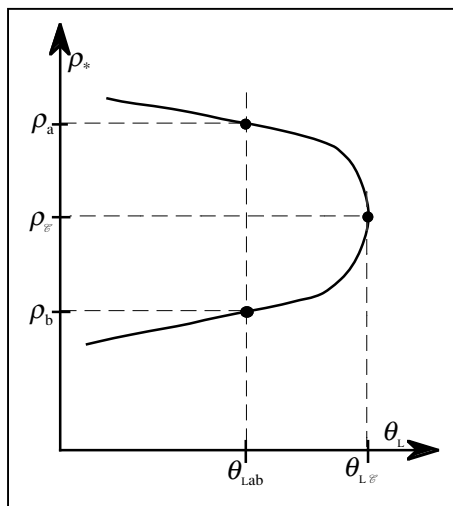


Figure 5-24.

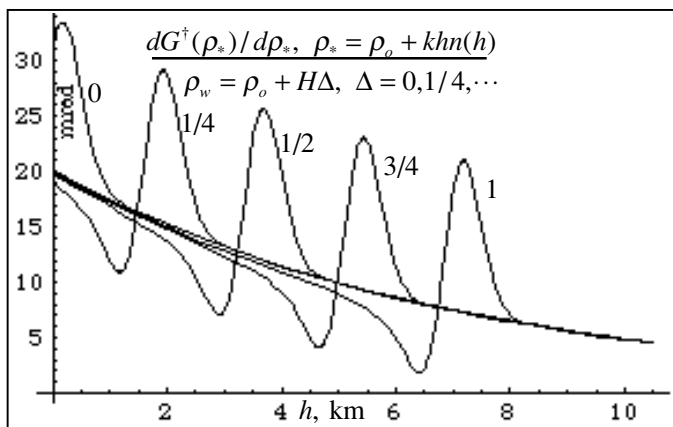


Figure 5-25.

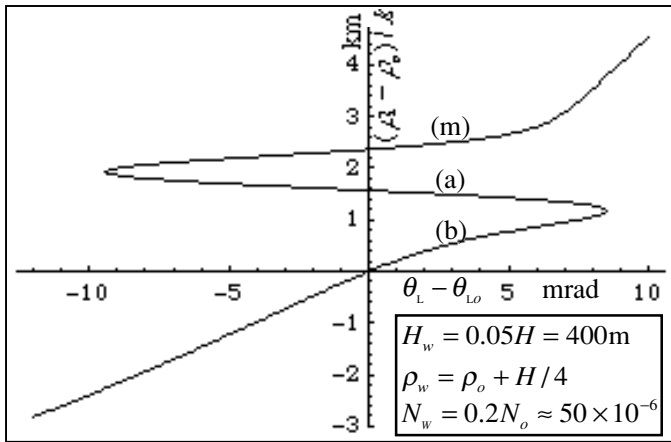


Figure 5-26.

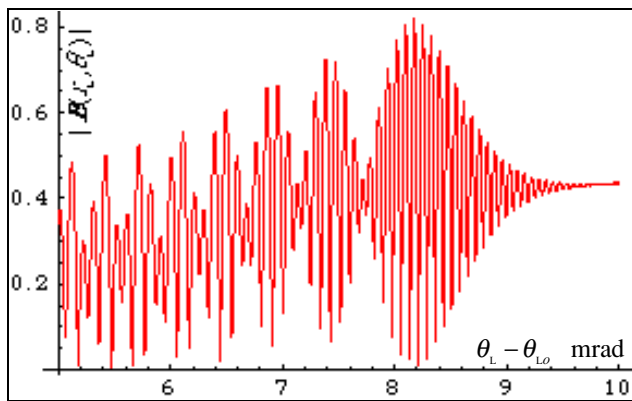


Figure 5-27.

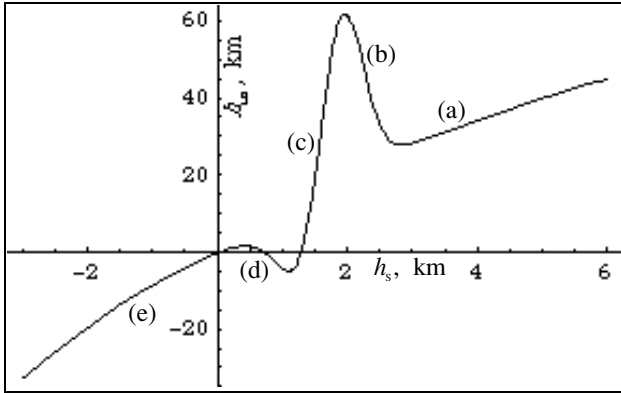


Figure 5-28.

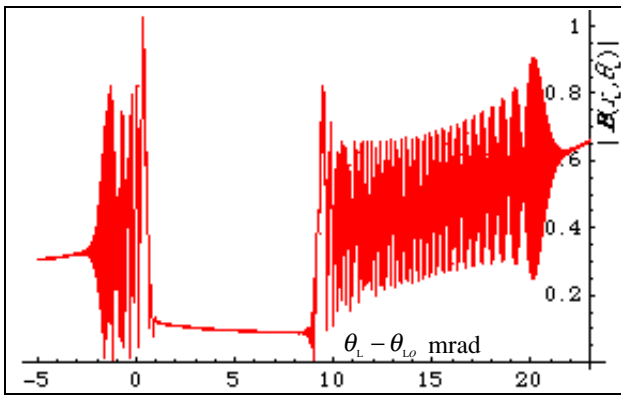


Figure 5-29.

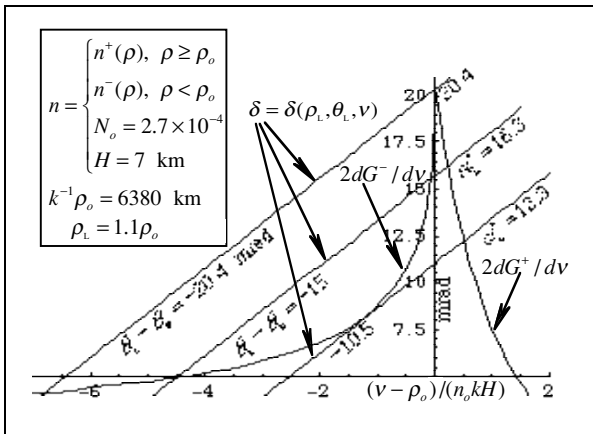


Figure 5-30.

Figure 5-31.

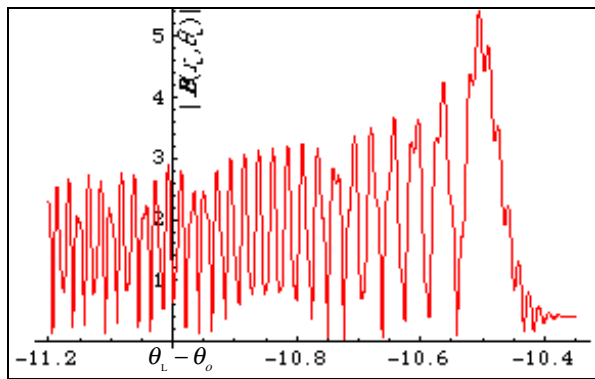
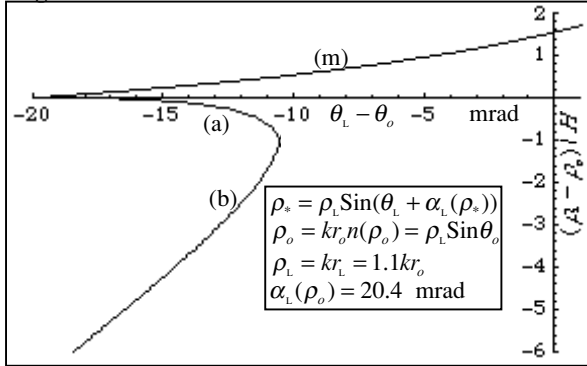


Figure 5-32.

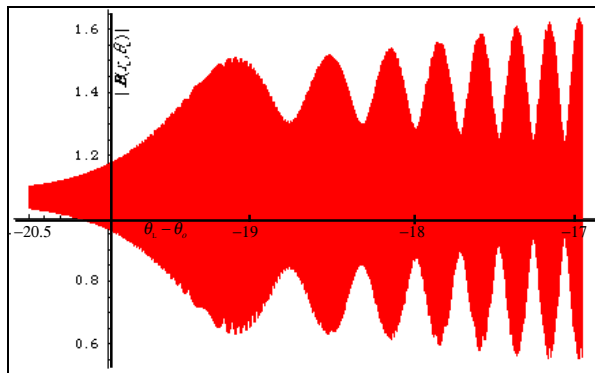


Figure 5-33.

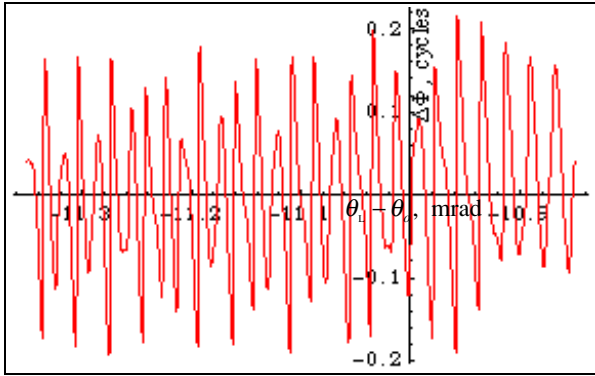


Figure 5-34

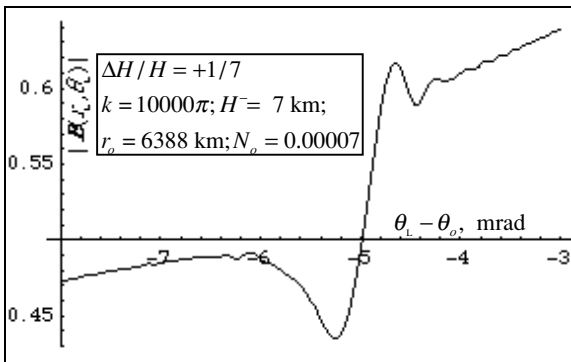


Figure 5-35.

Figure 5-36.

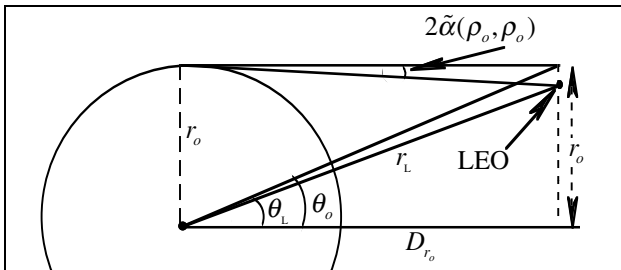
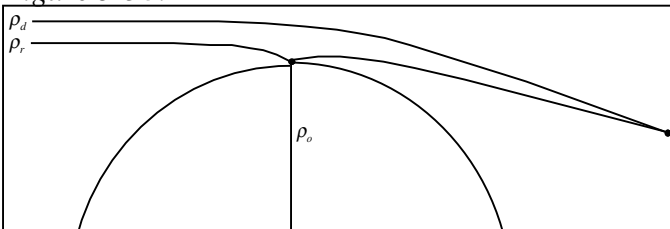


Figure 5-37.

Figure 5-38.

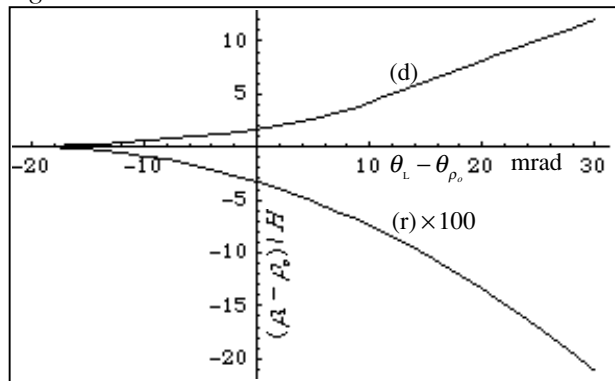


Figure 5-39

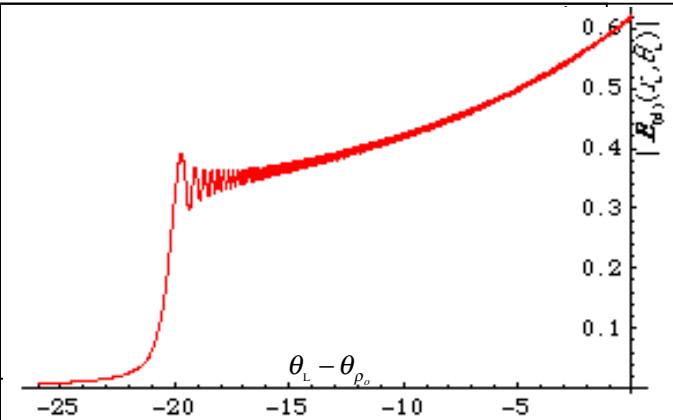
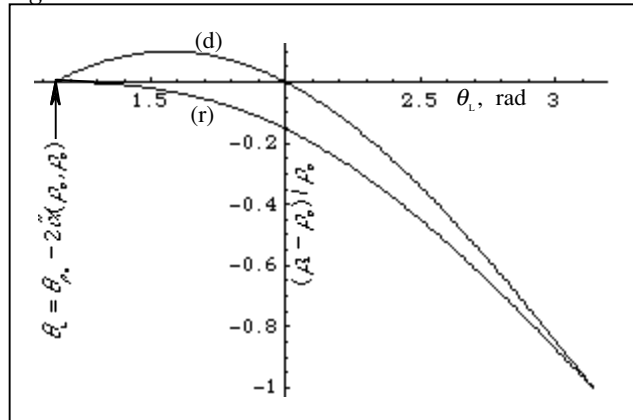


Figure 5-40.

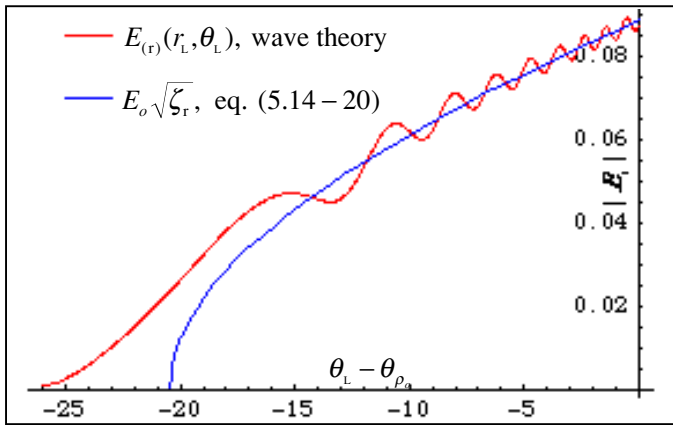


Figure 5-41.

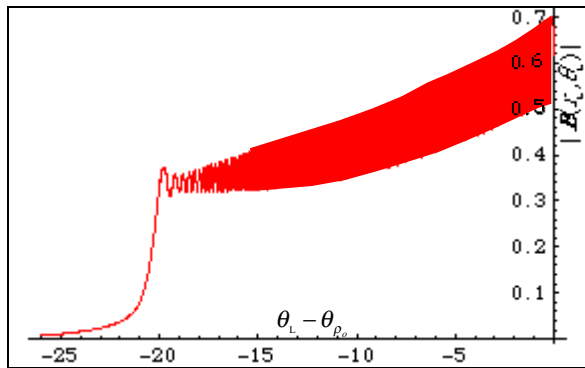


Figure 5-42.

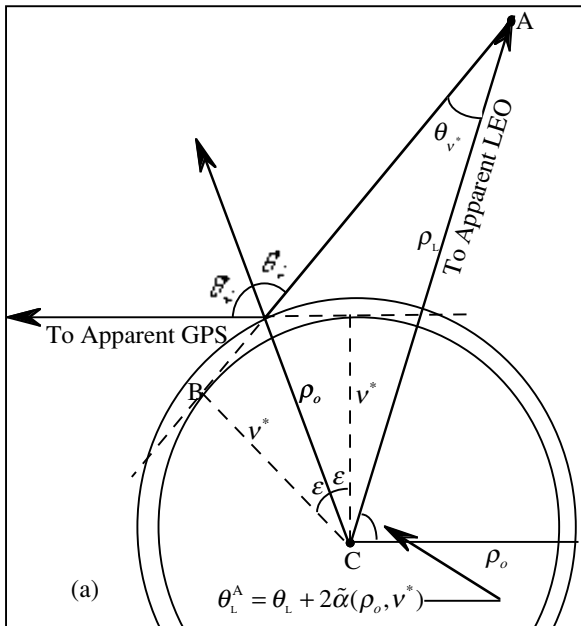


Figure 5-43a.

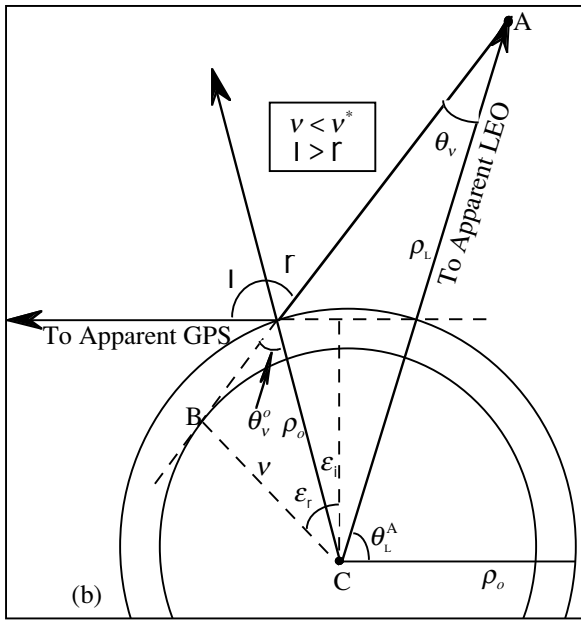


Figure 5-43b

Chapter 6

The Inverse Problem: Using Spectral Theory to Recover the Atmospheric Refractivity Profile

6.1 Introduction

Throughout Chapter 5 we mostly were concerned with the use of a full spectral representation of the electromagnetic field. We are not unmindful, however, of the potential suitability of this particular spectral technique to the inverse problem: determining the bending angle and refractivity profiles from a time-sequence of measurements of amplitude and phase of the received signal. This is especially interesting when adverse signal conditions prevail, that is, when the received signal exhibits significant amplitude and phase interference from multiple rays. In these situations classical recovery algorithms, such as those using the excess Doppler with the Abel transform, can run into difficulty because of the non-uniqueness or even non-existence of the ray path. Caustic points also lead to a breakdown in the validity of geometric optics, i.e., second order ray theory, on which the Abel transform algorithm is based. Spectral techniques, whether they rooted in geometric optics or in a full spectrum wave theory, can deal with these types of propagation problems. Although multiple rays with different bending angles can arrive at the LEO at the same time, they can not arrive simultaneously with the same excess Doppler values when spherical symmetry applies; those must be distinct. The transformation of the time series of observations into a spectral series provides a means to uniquely recover bending angle and refractivity profiles because of this one-to-one relationship between bending angle and excess Doppler.

In this chapter we briefly outline the use of the particular full spectrum wave theory technique developed in Chapter 5 for the inverse problem. The

spectral density function $G[\rho, \nu]$, which is the extra phase delay at the radial position ρ induced by the refractivity gradient on the ν th spectral component of the wave, plays a central role. We will not discuss in detail the computational aspects using this particular spectral technique, nor the concomitant use of estimation theory techniques on noisy data.

We assume that the LEO has received an ordered time series of amplitude and phase measurements, appropriately stripped of the geometric LEO/GPS Doppler signature and any other contributory error sources. For example, we assume perfect orbit knowledge from POD. We assume that the effects of oscillator variations in the various clocks affecting the phase measurements have been eliminated through redundant differential tracking; and we assume that ionospheric effects can be eliminated through dualband L1 and L2 tracking, possibly supplemented with modeling to eliminate third order effects. We also assume that the SNR of the signal is sufficient so that a Nyquist sample rate is practicable with respect to the bandwidth of the particular atmospheric signature under study.

Although this chapter does discuss a stand-alone Fourier approach for the inverse problem with radio occultation data, we note again that a principal utility of radio occultation data is for meteorology and numerical weather prediction. There the occultation data are merged in a timely way into a much richer and broader multi-sensor data base, which is constrained by a comprehensive model that characterizes the atmosphere and controls atmospheric processes. This is already mentioned in Chapter 1. In this context the difference between an actual radio occultation observation and a predicted observation obtained by forward propagation of the GPS signal through the atmospheric model becomes a constraint among the free parameters of the model. The 4DVAR methodology for minimizing a quadratic cost function involving disparate competing data bases constrained by the model is central to accurate weather prediction [1]. In this application one usually would not recover the refractivity profile from a radio occultation profile alone, but only from within the 4DVAR context. Nevertheless, the stand-alone approach has some important uses, and it is discussed in the next sections.

6.2 GPS Receiver Operations

It is helpful to understand some basic operational aspects of the GPS receiver in measuring phase and amplitude. The BlackJack GPS receiver series is a modern, high accuracy, dualband, digital receiver developed by JPL for scientific applications in space. As its development has evolved over the past several years it has successfully flown on over 1/2 dozen Earth satellites for navigation and time keeping, precision orbit determination, geopotential mapping, ocean reflections and limb sounding [2].

The following is a rather high level account of this particular receiver. It omits important details, which are crucial to the actual recovery of atmospheric parameters, but which are less relevant to the discussion here, e.g., dual frequency tracking to remove ionosphere effects, dealing with clock errors, digital signal processing and correlation operations on noisy data [3,4].

In normal operations the BlackJack receiver uses a closed-loop phase model before extracting the phase delay information to reduce the frequency of the received RF signal from the GPS satellite to baseband, a few tens of Hertz. First, the received RF signal (carriers in the 1.2–1.6 GHz range) is 1-bit digital sampled in-phase and separately in-quadrature, that is, with the received RF phase shifted by 90 deg. To facilitate the signal processing by the receiver, the data rate of these two parallel bit streams is reduced to an intermediate frequency (IF) of around 200 KHz. The IF frequency of these bit streams is then further reduced to the baseband frequency using an in-receiver phase model. This model is generated from a cubic polynomial fit to previous phase measurements made by the receiver over the past few tens of milliseconds, and in the closed-loop mode it is updated after each observational epoch with the latest phase measurement [Chk w. Meehan xx].

The L1 (1575 MHz) and L2 (1226 MHz) carriers of the navigation signals from each GPS satellite are derived from the same onboard master oscillator, and therefore, they are initially coherent. The L1 carrier is phase-modulated coherently with the C/A and P ranging codes. The L2 carrier is phase-modulated only with the P-code, but coherently with the L1. These ranging codes are pseudorandom, phase-modulating square waves that fully suppress the carrier tone. Each transition of a code, occurring at a frequency or chip rate of 1.023 MHz for the C/A code and at 10.23 MHz for the P code, involves a change in phase of the carrier of either zero or 180 deg in accordance with the pseudorandom algorithm specific to that particular code. Both carriers also are phase-modulated with a header code operated at a 50 Hz chip rate. This very low rate code carries the almanac and timing information for the tracked satellite and other satellite health and housekeeping data. Each satellite broadcasts distinctive codes that are unique to that satellite. These codes are mutually orthogonal and they also are orthogonal between GPS satellites. In the limit, cross-correlating two different codes yields a null result. Moreover, because the codes are pseudorandom, their auto-correlation function is triangular across an alignment offset of up to ± 1 chip period, and its value is zero outside of this range. This means that multiplying the same code with itself but time-shifted also yields a null result unless the two components are aligned within ± 1 chip period, the spatial equivalent of about ± 30 m for the P-code. Therefore, cross-correlating the received signal from a specific GPS satellite with the appropriately time and Doppler-shifted C/A and P-code replicas effectively filters out the signals received from all other satellites and also signals with phase delays greater than the chip period. Also, by aligning the

codes it enables a determination of the propagation delay between that satellite and the receiver plus any time-keeping difference between their clocks.

The baseband signal embedded in the in-phase and the quadrature bit streams being processed by the receiver is a composite of all broadcasting GPS satellites in view of the receiver's antennas. Also, the sign of each bit, +1 or -1, is dominated by the inevitable thermal or SNR noise on the original RF sampling process. Each of these bi-level bit streams is multiplied on a bit-by-bit basis by the time-delayed and Doppler-shifted baseband replica of the actual signal broadcast from a specific GPS satellite. These products are then averaged over a given time interval, i.e., they are cross-correlated. The nominal averaging time for this cross-correlation is 20 milliseconds. The receiver can perform this operation in parallel on the dualband L1 and L2 signals from up to a dozen individual GPS satellites. The cross-correlation operation accomplishes three objectives. First, the signals from all other satellites are essentially filtered out. Second, when the time-delayed and Doppler-shifted baseband replica is aligned with the arriving signal from that specific satellite, the effective SNR noise on the correlation products for that satellite is averaged down. Third, the alignment returning the maximum value of the correlation product provides the difference of the reception epoch of the receiver clock minus the transmission epoch of the clock onboard the GPS satellite. Synchronizing the transmitter and receiver clocks then allows the propagation delay, or group delay, from that satellite to the receiver to be obtained. Clock synchronization is achieved through redundant concurrent tracking of selected GPS satellites, including the occulted satellite, from the LEO receiver and from ground station receivers. Concurrent tracking of multiple GPS satellites from multiple receivers allows determination of the clock epoch differences. This cross-correlation between the received signal and its replica from a given satellite is performed on the in-phase bit stream and also separately on the quadrature bit stream. Thus, on each bit stream the signal is averaged over 20,000 transition points or chip periods of the C/A code and 200,000 chip periods of the P code. The noise error in the average is inversely proportional to the square root of the averaging interval.

The propagation delay between the broadcasting GPS satellite and the receiver may be determined in two ways. The alignment of the bit streams, so that the autocorrelation function described above from the received and replica codes is maximized, yields the group delay plus SNR error. After the autocorrelation function is maximized, a far more precise determination of propagation delay is obtained from the measurement of the phase of the carrier itself, which has a wavelength of only about 20 cm; in comparison, the "wavelength" of the P-code is about 30 m. This propagation delay of the carrier phase is obtained from the two time-averaged correlation coefficients, the in-phase coefficient I and the quadrature coefficient Q. The arctangent of I divided by Q gives a measurement of the difference in true phase minus the

phase predicted from the in-receiver model, modulo 2π ¹. If no cycles are lost between the epochs of the current measurement and the immediate previous measurement 20 msec earlier, adding the predicted phase from the model to the measured phase difference gives the true phase plus the averaged-down SNR error at that observation epoch. An important property of this scheme (adding the predicted phase to the measured difference) is that the SNR measurement error of the true phase at each observation epoch is statistically independent of the SNR error at any other epoch. The root-sum-square of the in-phase and quadrature correlation coefficients gives the amplitude of the signal. The SNR error on the phase measurement is statistically uncorrelated with the SNR error on the amplitude measurement. These measurements of true phase and amplitude are then reported by the receiver at a sample rate of nominally 20 ms; this is a convenient reporting rate because of the 50 Hz header code. Other multiples of that canonical sample interval can be used.

This closed-loop correlation and reporting scheme for the phase works well when the RMS difference between the predicted and measured phase is small, substantially less than 1/4 cycle. In this case there is a very high probability that no complete cycles, of either a positive or negative integer number, have been unaccounted for between the two successive measurement epochs. Using Gaussian statistics it is easy to show why this is so. Suppose that the measured phase at a particular epoch differed from the predicted value by exactly 1/2 cycle. Then we would have no way of determining whether that measurement was a cycle above or a cycle below the predicted value; all cycles look alike. Regarding this measured 1/2 cycle difference, do we add it to or subtract it from the predicted phase given by the model? Suppose that the statistical difference between measured and predicted phase, arising from either SNR errors or from unknown phase acceleration or from both, turns out to be 1/4 cycle, $1 - \sigma$. Then the probability of getting less than 1/2 cycle difference at each measurement epoch, if the errors are Gaussian distributed, is $\text{Erf}[2 / 2^{1/2}]$, or 0.9545. It follows that the probability of having at least one difference that is greater than 1/2 cycle after n successive statistically independent samples is

¹ In early versions of the Blackjack receiver, a simpler 2-quadrant arctangent routine without complete 4-quadrant resolution was used on the I and Q correlation coefficients to extract phase. This design choice eliminated the requirement to determine during the signal processing the sign of the 50 Hz header code bit ± 1 . An error in sign affects the sign of the I and Q correlation products the same way, and cancels in their ratio. But, this results in a 1/2-cycle ambiguity. This exacerbates the cycle slipping problem in noisy and/or loop-stressed conditions. Unfortunately, this arctangent operation has resided in the digital signal processing firmware (the ASIC) of the receiver, which is not easily modified without incurring the expense of redesigning and fabricating the ASIC at a silicon foundry. BlackJack designers at JPL plan to incorporate a 4-quadrant discriminator in a future version. The discussion here assumes the 4-quadrant version.

$1 - 0.9545^n$. After 1 s of elapsed time at a rate of 1 sample per 20 ms, $n = 50$; therefore, for $\sigma = 1/4$ cycle, the probability of getting at least one phase difference greater than $1/2$ cycle in 50 trials is virtually certain. If at a given measurement epoch the measured phase difference is $1/2$ cycle, then there is at least a 50% chance of choosing the wrong integer for the cumulative cycle count. Therefore, for $\sigma = 1/4$ cycle, a non-zero integer number of cycle slips becomes virtually certain within 1 s with a 20 ms sample interval. On the other hand, suppose the $1 - \sigma$ value of the statistical difference between the measured and predicted phase is $1/8$ cycle. Then the probability of getting one or more phase differences that are greater than $1/2$ cycle in n successive samples is $1 - \text{Erf}\left[4 / 2^{1/2}\right]^n = 1 - 0.9968^n$, still nearly zero for $n = 50$. In good signal conditions with small differences between measured and predicted phase, this scheme (of adding the predicted phase to the measured difference modulo 2π) usually works well.

6.2.1 Adverse Signal Conditions

The problem arises in adverse signal conditions, for example, when more than one ray arrives concurrently at the LEO, leading to interference and to sharp accelerations in phase across the troughs in the amplitude scintillation. Another adverse condition is where the LEO enters a quasi-shadow zone where no rays or at most highly defocused rays are present. Interference scenarios have been described earlier in Figs. 5-19 to 5-34. Figure 5-34 shows a specific example of fringe frequencies of roughly 50 Hz, 1 cycle change over 20 ms. Reducing the refractivity gradient in this model by a factor of four to obtain more realistic conditions, that is, that are more closely aligned with the Earth's atmosphere, quadruples the length of time. Thus, $1/4$ cycle change occurs for this relaxed case in 20 ms, which still a potential problem for closed-loop operations. This level of phase acceleration is serious enough to cause with some probability on each correlation interval at least one cycle not to be properly added or subtracted from the integer count book-kept by the in-receiver phase model. Over the many successive 20 ms sample intervals spanning a few seconds this could amount to a significant number of systematically lost cycles.

Consider the Gaussian refractivity model used in Fig. 5-20 and also the composite Gaussian/exponential model in Fig. 5-26. For a Gaussian distribution the impact parameter separations between rays in the multipath zone scale roughly linearly with the $1 - \sigma$ width H_w of the distribution, but the width of the multipath zone (in θ_L or in elapsed time) scales roughly as $H_w^{-3/2}$. The separations between rays scale only weakly with N_w , but the width of the zone scales nearly linearly with N_w . For the impact parameter diagram shown in Fig. 5-20 the values used are $N_w = 0.0001$ and $H_w = 1.6$ km, which results in

impact parameter separations of roughly 10 km, rather large compared to realistic scenarios. But, in Fig. 5-26 the values used in the Gaussian component are $N_w = 0.00005$ and $H_w = 350$ m, corresponding to a peak water vapor number density of about 1% of the local dry air density. More importantly, for Fig. 5-26 this component is superimposed on a background refracting medium that is defocusing, an exponential profile corresponding to dry air. This composite model leads to impact parameter separations in the multipath zone of up to 3 km, or to excess Doppler differences of up to about 15 Hz, or 1/3 cycle in 20 ms.

Another adverse signal situation occurs when the LEO encounters a shadow zone with weak signal conditions followed by flaring and strong interference. Consider a local refracting medium (e.g., a water vapor layer) embedded in an ambient medium (e.g., dry air) that gives rise to the transient in Fig. 6-1, which shows bending angle versus impact parameter. An abrupt increase in refractivity below a spherical boundary could yield this form for the bending angle profile. Here $\rho_o = k(1 + N_o)r_o$ corresponds to sea level. N_o is the refractivity for dry air at sea level. The exponential model in Eq. (5.8-2a) is used here for the dry air component of the refractivity with $N_o = 0.00027$ and the scale height $Hk^{-1} = 7$ km. Below the boundary at $r_o + Hk^{-1}/4$, about 2 km above sea level, the total refractivity abruptly increases, but then with decreasing altitude its gradient gradually approaches the dry air gradient. Figure 6-7 shows the recovered refractivity profile for this water vapor layer.

The impact parameter diagram shown in Fig. 6-2 results from the same bending angle profile shown in Fig. 6-1. The values of the refractivity and orbit parameters are the same for Figs. 6-1 through 6-7. The LEO orbit radius is $r_L = 1.1r_o$. This figure, expressed in terms of impact parameter altitude in kilometers versus orbit angle, provides an example of a shadow zone (where $d\rho_*/d\theta_L \approx 0$) followed by caustic flaring and multipath. At $r \approx r_o + H/4k$ the main ray (m) encounters a sharp increase in refractivity as its tangency point descends below a boundary there. This causes a shadow zone. As θ_L further decreases the first caustic is encountered at $\theta_L - \theta_{\rho_o} \approx 2$, leading to the creation of two more rays (a) and (b) in addition to the main ray (m). For a coplanar geometry the Doppler difference between these new rays and the (m) ray is 6–7 Hz at the first contact point; it gradually increases as the rays separate. This difference is given by $\Delta f_{b-m} = (\rho_{*m} - \rho_{*b})\dot{\theta}_L / 2\pi$, where $\dot{\theta}_L$ is the LEO angular velocity in the plane of incidence. Interference continues until the (a) and (m) rays disappear below the lower caustic point at $\theta_L - \theta_{\rho_o} \approx -7$. This scenario has been discussed in more detail in Chapter 2 using a thin phase screen and scalar diffraction theory. In particular, Chapter 2, Figs. 2-2c and 2-12. In the case shown here in Fig. 6-2, the (a) and (b) rays created at the right-hand caustic

point have phases at the LEO that differ by many cycles from the phase of the main ray (m). Moreover, they also have substantially lower excess Doppler frequencies. In this figure the difference is 6–7 Hz initially for a coplanar geometry, and it grows to 15 Hz as the altitude of the (b) ray impact parameter separates from the other two. When the occulted GPS satellite does not lie in LEO orbit plane obliquity effects reduce this Doppler level by up to about 30%; see Eq. (6.3-4) and Fig. 6-8.

Figure 6-3 shows the amplitude of the field at the LEO that results from this transient in bending angle shown in Fig. 6-1. In this example the upper caustic (right) yields strong signal flaring, but flaring from the lower caustic (left) is more muted. The voltage SNR well into the shadow zone is about 11 dB below voltage SNR for GPS signals in a vacuum. Diffraction creates edge fringes, but it also softens rough edges, resulting in the more gradual decay of SNR at the beginning of the shadow boundary.

Figure 6-4 shows the phase difference between the complete field and the field from the main ray (m) only. This figure begins near the end of the shadow zone and includes the contact with the first caustic. For a coplanar geometry this figure covers about 3–4 s of elapsed time. The excess Doppler for the (m) ray becomes nearly constant at entry into the shadow zone. In fact, a ray with nearly constant excess Doppler is a very defocused ray (see Eq. (6.3-11)). In the lower troposphere defocusing from the dry air refractivity gradient compresses the wider altitude differences of these impact parameters. Nevertheless, this figure shows an abrupt change in Doppler, within a 20 to 30 msec interval, from zero to about 7 Hz for a coplanar geometry.

Figure 6-5 shows a blow-up of the amplitude and phase of the field at the LEO in Figs. 6-3 and 6-4 in the vicinity of $\theta_L - \theta_{\rho_o} = 2.1$ around a very deep amplitude trough in the interference fringes. Here the emerging (a) and (b) rays, still essentially coherent and at a point slightly earlier than the geometric optics prediction of the caustic contact point, have strengthened so that their combined amplitude at the point where their phase is opposite the phase of the (m) ray nearly matches its amplitude. This causes a near complete cancellation of the total field and a short burst of rapid phase acceleration. The solid dots in the figure denote measurement epochs on 20 ms centers (for $d\tilde{\theta}_L / dt = -1$ mrad/s). The lighter dots are predicted phase values at future epochs. The SNR error bars are relative; their actual values also depend on signal gain and processing technique. In these situations the RMS disparity between predicted and measured phase can exceed 1/4 cycle primarily because of the inability of the in-receiver phase model to anticipate adequately the phase acceleration from interference between these multiple rays, and because of poor SNR noise in the phase measurement. Reducing the sample time to mitigate phase acceleration effects and to catch stray cycles, adversely impacts the effective SNR of the sampled measurement, which further exacerbates the potential disparity.

Multipath situations like these have led the BlackJack designers to experiment with alternate sampling and reporting schemes, such as flywheeling the receiver. An additional technique is to use a realistic excess Doppler algorithm combined with the receiver fractional phase measurement to connect phase. Other open loop schemes are feasible; the receiver uses a realistic in-receiver model and simply reports multiple time-lagged correlation coefficients in an open-loop mode. This approach leaves it to the data analysts to recover the various phases from the multiple rays relative to the realistic model and also to extract their respective amplitudes.

6.2.2 Flywheeling

Flywheeling uses an extrapolation from the in-receiver phase model, set at an earlier time where the phase was considered to be unambiguous, to predict the phase at a future epoch, without updating the model with new information from the most recent epoch because of its uncertainty. Both the closed-loop and the flywheeling modes are depicted in Fig. 6-5. To simplify the sketch only a second order loop is shown, which is indicated by the sloped straight lines. But the receiver actually uses a third order technique to predict forward; the extrapolation lines in this figure should be curved. In this simplest form of the in-receiver model, the phase measurements from the previous two measurement epochs are used to set the Doppler and the phase. In Fig. 6-5 these two points are located at $\theta_L - \theta_{\rho_o} = 2.14$ and 2.16 mrad. In this figure time evolves to the left for a setting occultation at a rate of roughly -1s/mrad for a coplanar occultation. Thus, the abscissa can be read directly in seconds of time. For a non-coplanar case the time scale would be compressed by the obliquity; for a 30° LEO orbit plane inclination relative to the plane of incidence, the elapsed time interval in this figure would be 30% greater (see Fig. 6-8). In the closed-loop mode the receiver uses the measured phase at these two earlier epochs at $\theta_L - \theta_{\rho_o} = 2.14$ and 2.16 to predict the phase at the next epoch at $\theta_L - \theta_{\rho_o} = 2.12$. Upon obtaining the new phase measurement at this later epoch it then updates the Doppler and phase of the model to predict the phase at $\theta_L - \theta_{\rho_o} = 2.10$. In the flywheeling mode the receiver does not update the model at $\theta_L - \theta_{\rho_o} = 2.12$, and it extrapolates the model fixed at the earlier epochs to predict the phase at $\theta_L - \theta_{\rho_o} = 2.10$.

Although this example of near-complete extinction of the field may seem pathological, it happens. The sharp phase acceleration combined with increased SNR error in the measurement is problematical for the receiver. To show this, redraw the phase prediction lines in Fig. 6-5, to run from the opposite ends of the $1-\sigma$ phase measurement error bars at $\theta_L - \theta_{\rho_o} = 2.14$ and 2.12 for the closed-loop, and at 2.16 and 2.14 for the flywheeling mode. The worst case

combination for this example runs from the upper end of the error bar for the earlier point to the lower end of the error bar for the later point. For these combinations the new phase predictions at $\theta_L - \theta_{\rho_o} = 2.10$ for both the closed-loop and the flywheeling modes are more than 1/2 cycle different from the actual phase at this point. Both the closed-loop and flywheeling modes have difficulty with this example.

The receiver automatically changes to the flywheeling mode according to certain pre-set signal conditions related to SNR and measured phase residuals, typically around $\text{SNR}_v \approx 30$, and it can revert back to closed-loop tracking according to another set of conditions. But, as many of the figures in this monograph suggest, tracking in either the close-loop or the flywheeling modes can result in cycle losses. If, for example, the earlier Doppler from the still strong and unique (m) ray is used in a flywheeling mode to carry the struggling receiver through the later poorer SNR conditions across a shadow zone, what happens at the contact point with the upper caustic? It depends on the strength of that caustic. A weak caustic (see Fig. 5-21 (b)) reveals itself at the LEO as the envelope defining the amplitude of relatively high frequency interference fringes. There the main ray (m) is still dominant, and the frequency of the interference fringes depends on the difference in altitudes of the impact parameter of the (m) ray and the impact parameter of the caustic rays. The receiver may have difficulty tracking either one of these nascent (a) and (b) rays because of their continuing interference with the (m) ray. Figure 5-22 shows the phase acceleration spikes that result when the rays become comparable in amplitude.

On the other hand, in the examples given in Fig. 5-32 and in Fig. 6-3, the extraordinary signal strength at the caustic contact is likely to induce the GPS receiver to lock onto the phase of the field there. For strong caustics, the nascent (a) and (b) rays, which are temporarily coherent in their early stage, become the principal contributors to the field at the LEO. This is the case in Fig. 6-3 at the caustic contact near $\theta_L - \theta_{\rho_o} = 2.0$. Until the impact parameters of these two nascent rays have had time to separate after the caustic contact point, there is very little interference between them. The amplitude of the field can be very strong at the LEO, depending on the curvature of the impact parameter curve θ_L vs ρ_* at the caustic contact point (see Eq. (5.12-11)). The width of a caustic peak for strong nascent rays can be hundreds of milliseconds, roughly given by $\Delta t \sim 3 |d^2 \theta_L / d\rho_*^2|^{1/3} / \dot{\theta}_L$. This width can be several to many 20-ms correlation intervals. For examples, see Fig. 5-32 ($\Delta t \approx 150$ ms) and Fig. 6-3 ($\Delta t \approx 500$ ms); easy enough to be lured there by these strong and locally stable fields. In this case the receiver reverts from the flywheeling

mode² to closed-loop and takes on the Doppler of the new field formed by the composite of all the rays. In the case of Fig. 6-4, the new Doppler abruptly (<50 ms) becomes 7 Hz less than the original Doppler for the (m) ray. The (m) ray has not yet disappeared, but it is no longer “recognized”. One has a 7 Hz negative bias relative to the (m) ray resulting from the receiver jumping from one ray to the other, from (m) to (b). However, it should be noted in this example, especially if the receiver were in fact to remain in the closed-loop mode through this delicate transition shown in Fig. 6-5, that no cycles were lost. The reported phase by the receiver in this case would be the measured phase of the complete field at the LEO from all rays. The problem is with the interpretation of the measurements resulting from the implicit adoption of a single ray paradigm.

Probably a prudent rule of thumb declares that whenever a caustic-like feature is encountered in the amplitude data, this probably means that it is a strong caustic in order to stand out in noisy data. This feature then is a harbinger for subsequent multipath and for the possibility that certain rays may not be properly accounted for with the usual ray theory approach i.e., the Abel transform. In fact, the mere event of the receiver converting to the flywheeling mode, should serve as an alarm announcing multipath and that spectral techniques may be required.

6.2.3 Refractivity Error from a Single Ray Paradigm

Figure 6-6 shows the receiver in the flywheeling mode using the excess Doppler from the (m) ray just prior to entering the shadow zone ($\theta_L - \theta_{\rho_o} \approx 7$) to power through the shadow zone with poor SNR. Closed-loop operations resume with return of strong signal near the first caustic contact point ($\theta_L - \theta_{\rho_o} \approx 2$), a few seconds later. The continued existence of the (m) and (a) rays after the (b) ray is “tracked”, and their contributions to refractivity recovery, are essentially ignored in the multipath zone with the single ray paradigm. It is straightforward to calculate the error in recovered refractivity caused by ignoring this multipath.

² For the specific refractivity profile used in Figures 6-1 through 6-5, it is doubtful that a high performance receiver like the BlackJack actually would drop out of the closed-loop mode upon entering the shadow zone. Diffraction softens the rate of falloff in amplitude at the shadow boundary. Also the average minimum amplitude in this example is still 20% of the original amplitude. At $\theta_L - \theta_{\rho_o} \approx 2.1$ there is a single episode of almost complete cancellation between the (m) ray and the combined field from the emerging (a) and (b) rays. This event could force the receiver into the flywheeling mode, but it is somewhat irrelevant. In both modes, flywheeling or closed-loop, the resulting error comes more from the analysis of the tracking data than from the receiver. Following a single ray paradigm the error in the refractivity recovery would be essentially the same regardless of mode.

For example, one can adopt as the true refractivity model the same model used to generate Figs. 6-1 through 6-5. This includes an exponential model for dry air plus a localized Heaviside-like component for the water vapor layer. This causes the multipath zone shown in Fig. 6-2 with the three rays, (m), (a) and (b). The adopted model for recovery would be the exponential-only model, which has no additional Heaviside-like refractivity component and it allows no multipath. If we then extend the impact parameter curve from the (b) ray near the caustic up to the (m) ray, as shown in Fig. 6-6, we have essentially ignored the entire anomalous (a) ray and a segment of the (m) ray. These sections are shown as the dashed loop in this figure. A best fit of N_o and H from the exponential-only model to this region can be done. Comparison of the recovered refractivity profile from this fitted model with the true refractivity profile gives the error resulting from overlooking the dashed segments of the (a) and (m) rays. This is shown in Fig. 6-7. Here the true refractivity profile $N(\rho_*)$ generates the bending angle profile shown in Fig. 6-1, but the modeled refractivity profile $\hat{N}(\rho_*)$ is only an exponential fit to the observations. This shows the difference in refractivity profiles between the fitted model and the true model. In this case the error in the refractivity is negatively biased because the excess Doppler from the (b) ray, including its extrapolation backward, is on average less than the combination of the (a) and (m) rays. This systematic loss of counted cycles relative to the (m) and (a) rays is equivalent to a negative error in the Doppler and this effectively leads in this example to a negative error in the recovered refractivity profile. See Appendix F for a further discussion of the error in recovered refractivity expressed in terms of an error in excess Doppler or bending angle.

An examination of actual SNR profiles from LEO occultation observations reveals a seemingly endless number of scintillation episodes, particularly as the signal passes through the water vapor-laden lower troposphere or through various layers in the ionosphere. Flywheeling does not appear to provide a comprehensive remedy when deep scintillation occurs. It probably will be augmented by the JPL group soon in favor of some open-loop scheme, such as reporting correlation coefficients using a realistic Doppler model, and with multiple time-lags between the predicted and received time series.

6.3 Spectral Representation of the Field at the LEO

With these caveats, we now assume that an appropriate signal processing scheme has been implemented so that the amplitude and connected phase has been recovered from the receiver, plus the unavoidable SNR noise. We start from the spectral representation for an outgoing wave evaluated at the LEO located at (ρ_L, θ_L) . For the emitting GPS satellite located at (ρ_G, θ_G) with $\theta_G = \pi$, a fixed value, the spectral integral representation for the field is given

from Eqs. (5.9-5) and (5.9-6), modified to account for the finite value of ρ_G . We have

$$\left. \begin{aligned} E(\rho_L, \theta_L) &= \frac{E_o}{\sqrt{2\pi\rho_L \sin\theta_L}} \int_0^\infty \left(\frac{\sin\theta_v^L}{\cos\theta_v^G \cos\theta_v^L} \right)^{\frac{1}{2}} \exp(i\Psi(+, -)) dv, \\ \Psi(+, -) &= D_v^G + D_v^L + v(\theta_v^G + \theta_v^L - \theta_L) - 2G^\dagger(v) - \frac{\pi}{4}, \\ \theta_v^G &= \sin^{-1}\left(\frac{v}{\rho_G}\right), \quad \theta_v^L = \sin^{-1}\left(\frac{v}{\rho_L}\right), \quad v < \rho_L < \rho_G, \\ D_v^G &= \sqrt{\rho_G^2 - v^2}, \quad D_v^L = \sqrt{\rho_L^2 - v^2} \end{aligned} \right\} \quad (6.3-1)$$

where³

³ The spectral representation for the field at the LEO when the GPS satellite is located at a finite distance, about 4.5 Earth radii, must account for the wavefront curvature of the incident wave. This is a spherical wave of the form $e^{i\rho_{LG}}/\rho_{LG}$ centered at the emitting GPS satellite (see Figure A-3). Although we did not derive this form in Eq. (6.3-1), its extra terms compared with the collimated form given in Eqs. (5.9-5) and (5.9-6) can easily be inferred from the difference in the asymptotic forms at large distances out of the atmosphere for the incoming spectral coefficients $a_l^-(\rho)$ for these two cases. These are given in Eqs. (5.5-3a) for the case of a collimated or planar incident wave and in Eq. (5.5-3b) for the case of a spherical incident wave. For the latter, $a_l^-(\rho)$ carries the extra factor $i^{l+1}\xi_l^+(\rho_G)/\rho_G$, which is derived from the multipole spectral expansion for a spherical wave [5] combined with the addition theorem for spherical harmonic functions. But, for $\rho_G \gg v$, we may use the asymptotic form for the spherical Hankel function $\xi_l^+(\rho_G)$. This factor has the asymptotic form

$$i^{l+1}\xi_l^+(\rho_G) \rightarrow (\rho_G^2/(\rho_G^2 - v^2))^{1/4} \exp[i((\rho_G^2 - v^2)^{1/2} + v\theta_v^G)]$$

which coincides with the extra terms in Eq. (6.3-1). Here the phase delay spectral density function $\Psi(+, -)$ is referenced to the emitting GPS satellite, whereas $\Psi(+, -)$ in Eqs. (5.9-5) and (5.9-6) for the collimated case is referenced to the line $\theta = \pi/2$. Also, the constant E_o has a different meaning from E_o in Eq. (5.9-5). Essentially it must account for the $1/\rho_{LG}$ space loss that the amplitude of the spherical wave emitted from the GPS satellite incurs in traveling to the LEO. This is inconsequential in recovering the refractivity profile because it is the variability of the amplitude and phase over an occultation episode that contains the atmospheric information. The product $(\cos\theta_v^L \cos\theta_v^G)$ in Eq. (6.3.1) is related to the reduced limb distance, $D = D_G D_L / (D_G + D_L)$, with $D_G = \rho_G \cos\theta_{\rho^*}^G$ and $D_L = \rho_L \cos\theta_{\rho^*}^L$.

The ray theory interpretation of θ_v^L and θ_v^G is as follows. Let $v = v^*$ be a spectral number at which Ψ assumes a stationary value. When super-refracting situations are

$$\left. \begin{aligned}
 2G^\dagger(\nu) &\doteq 2 \int_{\rho^\dagger}^{\infty} \frac{d \log n}{d \rho} g(\hat{y}) d\rho \doteq - \int_{\nu}^{\infty} \alpha_L(\omega) d\omega, \\
 g(\hat{y}) &= \pi K_\nu^2 \left(\text{Ai}'[\hat{y}]^2 + \text{Bi}'[\hat{y}]^2 - \hat{y} \left(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \right) \right), \\
 \hat{y} &\doteq \frac{1}{4K_\nu^4} (\nu^2 - \rho^2), \quad K_x = \left(\frac{x}{2} \right)^{\frac{1}{3}}, \\
 \nu &= l + \frac{1}{2}, \quad \rho^\dagger = \nu - \hat{y}^\dagger K_{\rho^\dagger}, \quad \hat{y}^\dagger = 0.44133 \dots
 \end{aligned} \right\} \quad (6.3-2)$$

The spectral density function $G^\dagger(\nu) = G[\rho^\dagger(\nu), \nu]$ for the phase delay from the refractive gradient and $\rho^\dagger(\nu)$ are discussed more fully in Section 5.7, and their forms are given in Eqs. (5.7-2) and (5.7-11), respectively⁴. They are the spectral density forms applicable to an outgoing wave at the position of the LEO. We have taken minor license with Eq. (5.9-5) by consolidating the radial and transverse components into a single scalar form for the electric field $E(\rho_L, \theta_L)$.

The resulting error is negligible for occultations.

We note that the spectral density for the phase, $\Psi(+, -)$, is a function of $(\rho_G, \rho_L, \theta_L, \nu)$. $\Psi(+, -)$ is the appropriate spectral density function for a position located well into the upper half-plane, $\pi \gg \theta_L \gg 0$, and in the outgoing quadrant, $\pi/2 - \alpha_L \gg \theta_L \gg 0$. $\Psi(+, -)$ gives the spectral density of the complete phase delay at the LEO position (ρ_L, θ_L) relative to the emitting GPS satellite located at (ρ_G, θ_G) with $\theta_G = \pi$. It includes the geometric delay terms and the term $-2G^\dagger(\nu)$ for the delay from the refractive gradient. The geometric delay term, D_ν^L , gives the delay in phase along a straight line between the LEO and the tangency point of the line on a sphere of radius $\nu < \rho_L$ centered at the origin. The term $\nu(\theta_L - \theta_\nu^L)$ is an arc length along this sphere of radius ν and it is subtracted from D_ν^L to correct it to the intersection of the sphere with the line $\theta = \pi/2$, which is the fixed reference line for computing phase delays at the LEO. Similarly, the terms $D_\nu^G + \nu\theta_\nu^G$ give the

avoided, we know that to high accuracy $\nu^* \doteq \rho_*$, the impact parameter of the corresponding ray. Then from Eq. (6.3-1) and Bouguer's law it follows that $\theta_{\rho_*}^L$ becomes the angle between the ray path tangent vector and the radius vector of the LEO, $\chi_L + \delta_L$, in Figure A-3. Similarly, $\theta_{\rho_*}^G \rightarrow \chi_G + \delta_G$. Note from Figure A-3 with $\theta_G = \pi$, it follows that $\theta_{\rho_*}^G + \theta_{\rho_*}^L - \theta_L \rightarrow \delta_G + \delta_L = \alpha_L$.

⁴ Incidentally, the difference between $\rho^\dagger(\nu)$ and ν is very small for a large sphere, $r_o/\lambda \gg 1$. Here that spatial difference is about 7 m.

geometric phase delay from the emitting GPS satellite along the straight line to the tangency point on the sphere of radius ν and thence along the sphere to the line $\theta = \pi/2$. For the LEO located at a given point (ρ_L, θ_L) , the spectral neighborhoods around the stationary phase points ν^* , that is, where $\partial\Psi/\partial\nu|_{\nu^*} = 0$, provide the principal contributions to this spectral integral in Eq. (6.3-1).

$E(\rho_L, \theta_L)$ is the scalar field at the LEO; the phase of $E(\rho_L, \theta_L)$ includes an integer multiple of 2π driven by the absolute phase represented by the spectral density function $\Psi(+, -)$. But, as we just discussed, it is problematic whether or not the actual phase measurements can yield unambiguous connected phase at all times. Nevertheless, it is important to have the correct phase change between observation epochs, uncorrupted by systematic cycle slipping by the receiver or by the post-measurement data editing. Occasional cycle breaks, although undesirable, are probably inevitable. We assume here that cycle breaks have been fixed.

6.3.1 Stopped Field at the LEO.

Let $\hat{E}(\rho_L, \theta_L)$ be defined as the “stopped” or “counter-rotated” signal received by the LEO. In this case “stopped” means that the orbital Doppler tone between the LEO and the observed GPS satellite has been removed. Also, the excess Doppler based on a first order model for the atmospheric refractivity signature has been removed. Thus

$$\hat{E}(\rho_L, \theta_L) = E(\rho_L, \theta_L) \exp(-i\varphi_m(t)) \quad (6.3-3)$$

where the model phase $\varphi_m(t)$ is a known function that describes the time history of the stopping phase. Dealing with $\hat{E}(\rho_L, \theta_L)$ rather than $E(\rho_L, \theta_L)$ alleviates aliasing problems in finite sampling techniques and sharpens the resolution. To get an idea of the magnitude of the variability of $\varphi_m(t)$, we will form the time derivative of $\Psi(+, -)$, the spectral density for the phase delay at the LEO given in Eq. (6.3-1). We evaluate Ψ at a stationary point in spectral number $\nu^* \doteq \rho_*$ where $\partial\Psi/\partial\nu = \theta_\nu - \theta_L - 2dG^\dagger/d\nu = 0$, and then we differentiate it with respect to time to obtain $\dot{\Psi}_*$. This has already been discussed in Section 5.12 for a circular LEO orbit with a coplanar geometry, that is, with the GPS satellite located in the orbit plane of the LEO.

6.3.2 The Obliquity Factor.

However, we also should allow for the obliquity effect because in general coplanarity does not apply. From Fig. 6-8 we have two angles defining the angular position of the LEO, θ_L and $\tilde{\theta}_L$. Here θ_L gives the angular position in the LEO orbit plane, but $\tilde{\theta}_L$ gives it in the propagation plane, which is defined

by the ray from the GPS satellite located in the negative z direction⁵ to the LEO. This propagation plane includes the “geocenter” (the center for the local geoid) and defines the great circle arc AC on the unit sphere in Fig. 6-8. The LEO orbit plane defines the great circle arc BC. The departure from coplanarity is given by the inclination angle I of the LEO orbit plane about the x -axis in the figure relative to the direction to the GPS satellite. The inclination angle I is satellite position-dependent, but it is readily expressed in terms of the orbit elements for the two satellites. The boresight-offset angle is ε . This is the azimuthal angle about the radial axis relative to the in-orbit plane direction at which the GPS satellite would be seen from the LEO. The spherical triangle ABC in Fig. 6-8 gives the relationship between θ_L and $\tilde{\theta}_L$. We have

$$\left. \begin{aligned} \sin \varepsilon &= \frac{\sin I}{\sin \tilde{\theta}_L}, \quad \tan \tilde{\theta}_L = \tan \theta_L \cos \varepsilon, \\ \cos \tilde{\theta}_L &= \cos \theta_L \cos I, \\ \frac{d\tilde{\theta}_L}{d\theta_L} &= \cos^2 I \cos \varepsilon = \frac{\cos^2 I}{\sin \tilde{\theta}_L} \sqrt{\sin^2 \tilde{\theta}_L - \sin^2 I} \end{aligned} \right\} \quad (6.3-4)$$

This provides the relationships between θ_L and $\tilde{\theta}_L$, and also the obliquity factor $d\tilde{\theta}_L/d\theta_L$ to reduce the Doppler for the effect of non-coplanarity. It follows that

$$\dot{\tilde{\theta}}_L = \left(\frac{d\tilde{\theta}_L}{d\theta_L} \right) \dot{\theta}_L \quad (6.3-5)$$

This obliquity factor is essentially constant over an occultation episode; thus, the relationship between $\tilde{\theta}_L$ and θ_L is essentially linear. The obliquity factor is shown in Fig. 6-9 as a function I . Here $\tilde{\theta}_L = \sin^{-1}(r_o/r_L)$ with $r_L/r_o = 1.1$, which gives about the correct value of $\tilde{\theta}_L$ during an occultation. Thus, an inclination of 30 deg reduces the excess Doppler to about 2/3 of the coplanar value. It is $\tilde{\theta}_L$ that should be used in Bouguer’s law and in the spectral density functions involving phase. Accordingly, we adopt the following convention in the subsequent discussion. In any expression that involves a sensitive variable, such as phase or angular velocity, we shall replace θ_L with $\tilde{\theta}_L$; otherwise, we leave the notation as is.

6.3.3 Doppler Variability.

⁵ We assume here for calculation of the obliquity factor that the emitting GPS satellite is located infinitely afar in the negative z -direction, $\theta_G = \pi$.

We now estimate the variability of $\varphi_m(t)$. For the purpose of calculating the Doppler variability we assume circular orbits for the satellites. From Eq. (5.12-4) it follows after this replacement of θ_L with $\tilde{\theta}_L$ in Bouguer's law that

$$\dot{\Psi}_* = -\rho_* \dot{\tilde{\theta}}_L \quad (6.3-6)$$

where the phase $\Psi_* = \Psi(\rho_G, \rho_L, \tilde{\theta}_L, \rho_*)$, and ρ_* is the impact parameter corresponding to a specific ray, not necessarily unique. To simplify this calculation we assume that the emitting GPS satellite is at an infinite distance. Its actual finite distance and orbital motion has a minor effect on our estimate here. From Bouguer's law in Eq. (5.6-5) for the GPS satellite at an infinite distance we have

$$\rho_* = \rho_L \sin(\tilde{\theta}_L + \alpha_L) = \rho_o + D_L \alpha_L + O[\alpha_L^2] \quad (6.3-7)$$

where $\rho_o = \rho_L \sin \tilde{\theta}_L$ and $D_L = \rho_L \cos(\tilde{\theta}_L + \alpha_L)$. The quantity D_L is close to the distance (or reduced distance) in phase units (i.e., distance $\times 2\pi / \lambda$) from the LEO to the Earth's limb, even with a modest cosine effect from orbit inclination angle I factored in. For the 20 cm wavelength of the GPS signal, $\rho_o \approx 2 \times 10^8$ at sea level and for a LEO orbit radius $r_L = 1.1r_o$, $D_L \approx 1 \times 10^8$. It follows that

$$\dot{\Psi}_* \doteq -\rho_o \dot{\tilde{\theta}}_L - (D_L \dot{\tilde{\theta}}_L) \alpha_L \quad (6.3-8)$$

The first term on the RHS of Eq. (6.3-8) produces the orbital Doppler term due to the LEO. For a typical LEO orbit $\dot{\tilde{\theta}}_L \approx 1$ mrad/s and it is essentially constant with time; thus, for the LEO part, $\rho_o \dot{\tilde{\theta}}_L / 2\pi \approx 30$ KHz times the obliquity factor $d\tilde{\theta}_L / d\theta_L$, which is shown in Fig. 6-9 as a function of the inclination angle I of the LEO orbit plane.

The orbital velocity of a GPS satellite is about half the LEO velocity because its orbit radius is about four times larger. During an occultation the position of a GPS satellite is located about 4.5 Earth radii away from the LEO on the far side of the Earth; so, only about a quarter of its velocity vector projects plus or minus in the direction of the LEO. Also, the GPS satellite orbit planes are inclined differently to the LEO orbit plane, and the limb of the Earth as seen from the LEO is offset downward from the LEO orbit velocity direction by roughly 25 deg. The upshot is that all of these factors combine to yield a maximum Doppler during an occultation from both LEO and GPS orbital kinematics of around 35 KHz.

The second term in Eq. (6.3-8) is the excess Doppler from the bending angle α_L caused by the atmospheric refractivity; here $D_L \dot{\tilde{\theta}}_L / 2\pi \approx 15(d\tilde{\theta}_L / d\theta_L) \text{ Hz/mrad}$. Dry air yields a bending angle for a ray path tangency point at sea level of about 20 mrad; so, this component reaches about 300 Hz. But, bending angles through water vapor in the lower troposphere can exceed twice this level. The water vapor contribution is largely unknown *a priori*, but it can be characterized statistically by latitude and season. Excess Doppler signatures typically range over several hundred Hz.

What about the variability of these Doppler frequencies? Let us assume that the LEO is in a circular orbit. Then from Eq. (6.3-6) it follows that

$$\ddot{\Psi}_* \doteq -\dot{\rho}_* \dot{\tilde{\theta}}_L \quad (6.3-9)$$

The term $\rho_* \ddot{\tilde{\theta}}_L \ll \dot{\rho}_* \dot{\tilde{\theta}}_L$ and it is ignored here. The acceleration contribution from the GPS satellite will be small over the relatively short time intervals of interest here. Upon differentiating ρ_* from Bouguer's law in Eq. (6.3-7), it follows that

$$\dot{\rho}_* = D_L \dot{\tilde{\theta}}_L \xi_L \quad (6.3-10)$$

where $\xi_L = (1 - D_L d\alpha_L / d\rho_*)^{-1}$ is the defocusing factor. It follows that the acceleration in the phase term is given by

$$\ddot{\Psi}_* \doteq -D_L \dot{\tilde{\theta}}_L^2 \xi_L \quad (6.3-11)$$

Similarly, it follows that the acceleration from the orbital motion of the LEO is given by

$$\frac{d^2 D_L}{dt^2} \doteq D_L \dot{\tilde{\theta}}_L^2 \quad (6.3-12)$$

Subtracting Eq. (6.3-12) from Eq. (6.3-11) gives the acceleration in excess phase

$$2\pi \dot{f}_D = \ddot{\Psi}_* - \frac{d^2 D_L}{dt^2} \doteq -D_L \dot{\tilde{\theta}}_L^2 \frac{d\alpha_L}{d\rho_*} \xi_L \xrightarrow{\alpha'_L \rightarrow -\infty} D_L \dot{\tilde{\theta}}_L^2 \quad (6.3-13)$$

It follows for strong defocusing that the excess Doppler rate approaches a constant value with time, 10-15 Hz/s, depending on the obliquity of the orbit and propagation planes. (Therefore, the bending angle rate of a given ray in

strong defocusing conditions will approach a constant, $d\alpha_L / dt \rightarrow \sim 1$ mrad/s, or equivalently, $d\alpha_L / d\tilde{\theta}_L \rightarrow \sim 1$)

Over the course of a few tens of seconds that we will be interested in recovering the refractivity profile under adverse signal conditions, how well can we fit the stopping phase with a simple linear polynomial in time $\varphi_m(t) = \omega_m t$? Eq. (6.3-13) suggests that in strong defocusing this linear form should be good. We set $\omega_m = \dot{\Psi}_*$. Then $\dot{\Psi} - \dot{\varphi}_m(t) \approx \ddot{\Psi}_* t$. From Eq. (6.3-11) it follows that $\ddot{\Psi}_* / 2\pi < \sim 2$ Hz/s in the lower troposphere where the defocusing factor from dry air has a value around 1/10. The defocusing in the lower troposphere causes the impact parameter separations between multipath rays to be proportionately compressed, thereby reducing their Doppler differences by a factor of about 10. Thus, for a 10s single-sided sample interval we can use a linear term in time to stop the Doppler in the signal with a frequency run-off of a few tens of Hz. Sampling the signal at a 50 Hz rate usually should satisfy the Nyquist criterion for this sample interval. At the altitude of the sporadic E layer the ambient value of the defocusing factor is essentially unity; we would need a narrower sample interval there.

The last question concerns the spread in Doppler tones from different multipath rays. We already have seen that the spread in altitudes of the tangency points typically is less than 5 km. Thus, the maximum spread in Doppler is less than $(D_L \dot{\theta}_L / 2\pi)(5k / D_L) \approx 25$ Hz; the bandwidth of most interference spectra is less than 10 Hz. A sample rate of 50 Hz should suffice.

Incidentally, Eq. (6.3-10), which gives the velocity of the impact parameter of a ray, shows the retardation caused by the defocusing factor ζ_L . In the lower troposphere the dry air component of the defocusing systematically compresses the altitude separation between multipath rays and narrows the maximum bandwidth of the interference spectrum, see Eq. (5.12-18).

6.4 Refractivity Recovery

Let us apply a discrete Fourier transform to the stopped LEO observations given in Eq. (6.3-3) and to the spectral representation for the stopped wave given in Eq. (6.3-1). For the latter we have

$$\left. \begin{aligned} \hat{E}[\omega] &= \sum_{k=-M/2}^{M/2} \hat{E}(\rho_L, \theta_L(t_k)) \exp\left(i\omega \frac{kT}{M}\right), \\ \tilde{\theta}_L(t_k) &= \tilde{\theta}_{L_o} + \dot{\tilde{\theta}}_L(t_k - t_o) = \tilde{\theta}_{L_o} + \dot{\tilde{\theta}}_L \frac{kT}{M}, \quad \theta_G \equiv \pi \end{aligned} \right\} \quad (6.4-1)$$

Here $\tilde{\theta}_{L_o} = \tilde{\theta}_L(t_o)$ is the LEO orbit angle at the center of the data interval measured in the plane of incidence, see Fig. 6-8. Eq. (6.3-4) provides the relationship between $\tilde{\theta}_L$ and θ_L . T is the total time span covered by the observations and $M+1$ is the total number (odd) of samples; for example, $M = 50T$ for a 50-Hz sample rate. Using the spectral representation in Eq. (6.3-1) for $\tilde{E}(\rho_L, \theta_L)$ and $\varphi_m(t) = \omega_m(t - t_o)$ for the stopping phase, we have from Eq. (6.4-1)

$$\left. \begin{aligned} \hat{E}[\omega] &= \sum_{k=-M/2}^{M/2} \hat{E}(\rho_L, \theta_L(t_k)) e^{i\omega \frac{kT}{M}} \\ &= \sum_{k=-M/2}^{M/2} e^{i(\omega - \omega_m) \frac{Tk}{M}} \left(\frac{E_o}{\sqrt{2\pi\rho_L \sin \tilde{\theta}_{L_o}}} \int_0^\infty \left(\frac{\sin \tilde{\theta}_v^L}{\cos \tilde{\theta}_v^G \cos \tilde{\theta}_v^L} \right)^{\frac{1}{2}} e^{i\Psi_L} dv \right) \\ &= \frac{E_o e^{i\Psi_{L_o}}}{\sqrt{2\pi\rho_L \sin \tilde{\theta}_{L_o}}} \int_0^\infty \left(\frac{\sin \tilde{\theta}_v^L}{\cos \tilde{\theta}_v^G \cos \tilde{\theta}_v^L} \right)^{\frac{1}{2}} \left(\sum_{k=-M/2}^{M/2} e^{ik(\omega - \omega_m - v\dot{\theta}_L) \frac{T}{M}} \right) dv, \\ \Psi_{L_o} &= D_v^G + D_v^L + v(\tilde{\theta}_v^G + \tilde{\theta}_v^L - \tilde{\theta}_{L_o}) - 2G^\dagger(v) - \frac{\pi}{4} \end{aligned} \right\} \quad (6.4-2)$$

Now we replace the discrete Fourier transform with the integral transform using the fact that

$$\sum_{k=-M/2}^{M/2} \exp(iuk) = \frac{\sin((M+1)u/2)}{\sin(u/2)} \xrightarrow{M \rightarrow \infty} 2\pi\delta(u), \quad -\pi \leq u \leq \pi \quad (6.4-3)$$

Here $\delta(u)$ is the Dirac delta function. Setting $u = \pi$ defines the Nyquist limit, $|\omega| \leq \pi M/T$. Upon replacing the discrete “delta function” in Eq. (6.4-2) with the Dirac delta function, it follows that the Fourier transform of $\tilde{E}(\rho_L, \theta_L)$ is given by

$$\left. \begin{aligned} \hat{E}[\omega_v] &= 2\pi E_o \sqrt{\frac{1}{2\pi\rho_L \sin \tilde{\theta}_{L_o}}} \left(\frac{\sin \tilde{\theta}_v^L}{\cos \tilde{\theta}_v^G \cos \tilde{\theta}_v^L} \right)^{\frac{1}{2}} \exp(i\Psi_{L_o}) \\ \Psi_{L_o} &= D_v^G + D_v^L + v(\tilde{\theta}_v^G + \tilde{\theta}_v^L - \tilde{\theta}_{L_o}) - 2G^\dagger(v) - \frac{\pi}{4} \\ \omega_v &= \omega_m + \dot{\theta}_L v, \quad D_v^G = \sqrt{\rho_G^2 - v^2}, \quad D_v^L = \sqrt{\rho_L^2 - v^2} \end{aligned} \right\} \quad (6.4-4)$$

For a finite sample interval the discrete Fourier transform yields the discrete delta function, which has the $\text{Sin}Nx/\text{Sin}x$ character shown in Eq. (6.4-3), and the actual time series would have additive noise. The granularity threshold from the discrete transform can be obtained from the first zero of the discrete delta function in Eq. (6.4-3), which occurs at $u = 2\pi/(M+1)$. This is equivalent to $\Delta\omega = 2\pi/T$, which is essentially the uncertainty inequality for a discrete process. In impact parameter space this relationship maps into a granularity in altitude of $\lambda\Delta\rho_*/2\pi = \lambda\dot{\tilde{\theta}}_L^{-1}\Delta\omega/2\pi = \lambda/(\dot{\tilde{\theta}}_L T) \approx 200/T \text{ m}$, or about 20 m for $T = 10 \text{ s}$. Instead of a Fourier transform, one could use any one of several other transform schemes, such as the Morlet wavelet transform, which treats the time series as a spectral composition of wave packets. The subsequent equations will differ and their efficiencies in recovering the refractivity profile might vary, but recovery should still be feasible. For the purpose of outlining this particular spectral approach for recovering the refractivity profile, we avoid further discussion of these important computational and related stochastic issues.

Note in Eq. (6.4-4) that $\omega_m \sim 2 \times 10^5 \text{ rad/s}$, or 30-35 KHz. On the other hand, $\omega_v/2\pi$ varies over only a few tens of Hertz within the time interval for which the Fourier transform is applied. We have used the slowly varying character of $\tilde{\theta}_v^L$ and $\tilde{\theta}_v^G$ to simplify the spectral expressions given in Eq. (6.4-4). Over the bandwidth spanned by ω_v , $\tilde{\theta}_v^L + \tilde{\theta}_v^G$ changes by less than 0.1%. Thus, we may set $\tilde{\theta}_v^L + \tilde{\theta}_v^G = \tilde{\theta}_{Lo} + \alpha_{Lo} \doteq \tilde{\theta}_{Lo}$ in the slowly varying terms (but not in Ψ_{Lo}). The error here is roughly 1% or smaller, the ratio $\alpha_L(t_k)/\tilde{\theta}_L(t_k)$. The Fourier transform in Eq. (6.4-4) further simplifies to

$$\begin{cases} \hat{E}[\omega_v] = 2\pi E_o \sqrt{\frac{1}{2\pi D}} \exp(i\Psi_{Lo}), \\ \Psi_{Lo} = D_v^G + D_v^L + \nu(\tilde{\theta}_v^G + \tilde{\theta}_v^L - \tilde{\theta}_{Lo}) - 2G^\dagger(\nu) - \frac{\pi}{4}, \\ \omega_v = \omega_m + \dot{\tilde{\theta}}_L \nu, \quad D^{-1} = D_L^{-1} + D_G^{-1} \end{cases} \quad (6.4-5)$$

We now take the Fourier transform over the occultation sequences of stopped phase and amplitude measurements made by the LEO given in Eq. (6.3-3). The temporal breadth of this sequence T would depend on one's goals for refractivity recovery. We equate these two Fourier transforms. The LHS on the upper line of Eq. (6.4-5) becomes the Fourier transform of the sampled amplitude and stopped phase of the field measured by the LEO; the RHS is from wave theory.

We note the one-to-one correspondence between ω_v or excess Doppler and the spectral number ν in wave theory, or the impact parameter in ray optics,

which holds when spherical symmetry applies. Also, we note that the bending angle may not be a unique function of time (or $\tilde{\theta}_L$), but it is a unique function of excess Doppler, and when spherical symmetry holds it is a unique function of impact parameter through Bouguer's law. The impact parameter diagrams in Figs. 5-20, 5-26, 5-28, 5-31 and 6-2 all show implicitly this uniqueness property of the bending angle versus impact parameter when spherical symmetry holds. See also Eqs. (6.3-6) and (6.3-8), which apply to a circular LEO orbit. We have converted through a Fourier transform the time-series of phase and amplitude measurements of the field, in which the bending angle may not be unique, into a spectral series in which the bending angle is unique (when spherical symmetry applies). It follows that we should be able to unambiguously determine the bending angle profile versus excess Doppler from the Fourier transform $\hat{E}[\omega]$, given, of course, the limitations imposed by measurement errors.

Returning to Eq. (6.4-5), the LHS is the Fourier transform $\hat{E}[\omega]$ from the observations. It is a known quantity. The RHS is from wave theory; it also contains quantities that are known *a priori* or from POD information, except for $G^\dagger(\nu) = G[\rho^\dagger(\nu), \nu]$. Forcing equality between these two Fourier transforms forces the phase of $\hat{E}[\omega]$ to equal Ψ_{Lo} . Therefore, we can determine values for $G^\dagger(\nu)$ from Eq. (6.4-5) over the Fourier bandwidth spanned by ω .

However, a more suitable platform for extracting values for $\log n$ is $dG^\dagger(\nu)/d\nu$. Recalling that $g(\hat{y}^\dagger) = 0$, it follows from Eq. (6.3-2) that

$$\frac{dG^\dagger}{d\nu} \doteq -\pi K_{\nu_\omega} \int_{\rho^\dagger}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho \quad (6.4-6)$$

Let us now differentiate the Fourier transforms in Eq. (6.4-5) with respect to ω . Noting that $d\omega_\nu / d\nu = \dot{\tilde{\theta}}_L$, we obtain

$$\frac{d \log \hat{E}[\omega_\nu]}{d\omega_\nu} = i \left(\tilde{\theta}_\nu - \tilde{\theta}_{Lo} - 2 \frac{dG^\dagger}{d\nu} \right) \dot{\tilde{\theta}}_L^{-1} \quad (6.4-7)$$

However, from Eq. (6.4-1) it follows that

$$\frac{d}{d\omega} (\hat{E}[\omega]) = i \int_{-\infty}^{\infty} t \hat{E}(\rho_L, \theta_L(t)) \exp(i\omega t) dt \quad (6.4-8)$$

No explicit differentiation of the observed phase and amplitude with respect to time is required to obtain the derivative of $\hat{E}[\omega]$ with respect to ω . In Eq. (6.4-6) we note that $\partial(\partial G[\rho, \nu] / \partial \nu) / \partial \rho \neq 0$ at $\rho = \rho^\dagger(\nu)$ (whereas

$\partial G[\rho, \nu] / \partial \rho = 0$). It follows that Eq. (6.4-6) provides a stable means for determining the profile for $d \log n / d \rho$ from the determination of $dG[\rho^\dagger(\nu), \nu] / d\nu$.

Let the Fourier observation function $F[\omega]$ be defined by

$$\dot{\tilde{\theta}}_L F[\omega_\nu] = i \dot{\tilde{\theta}}_L \frac{d \log \hat{E}[\omega_\nu]}{d\omega_\nu} + (\tilde{\theta}_{\nu_\omega} - \tilde{\theta}_{L_o}) \quad (6.4-9)$$

$F[\omega_\nu]$ is a determined spectral quantity from the measurements and the POD information. From Eqs. (6.4-6)-(6.4-11) it follows that

$$\left. \begin{aligned} -2\pi K_\nu \int_{\rho^\dagger}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho &= 2 \frac{dG^\dagger(\nu)}{d\nu} = \dot{\tilde{\theta}}_L F[\omega_\nu] \\ \rho^\dagger &= \nu - \hat{y}^\dagger K_\nu = \frac{\omega_\nu - \omega_m}{\dot{\tilde{\theta}}_L} - \hat{y}^\dagger K_\nu \end{aligned} \right\} \quad (6.4-10)$$

Eq. (6.4-10) is in effect a linear system, an integral equation from which the profile $d \log n / d \rho$ may be recovered from a spectral sequence of known values for $F[\omega]$ based on the spectral derivative of the Fourier transform of the observations. Recalling Eqs. (5.4-3) and (3.8-7), the negative argument asymptotic forms for the Airy functions are $\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \rightarrow \pi^{-1}(-\hat{y})^{-1/2}$ with $\hat{y} \doteq K_\nu^{-4}(\nu^2 - \rho^2) / 4$. Eq. (6.4-10) becomes

$$\left. \begin{aligned} -2\pi K_\nu \int_{\rho^\dagger}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho &\rightarrow \\ -2\nu \int_{\nu}^{\infty} \frac{d \log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - \nu^2}} &= \alpha_L(\nu) \end{aligned} \right\} \quad (6.4-11)$$

It follows from Eqs. (6.4-10) and (6.4-11) that

$$\alpha_L(\nu) \Leftrightarrow F[\omega_\nu] \dot{\tilde{\theta}}_L, \quad \omega_\nu = \omega_m + \dot{\tilde{\theta}}_L \nu \quad (6.4-12)$$

To the extent to which the asymptotic forms for the Airy functions are applicable, $F[\omega_\nu]$ is proportional to the bending angle for an impact parameter value of ν . It follows for a setting occultation ($\dot{\tilde{\theta}}_L < 0$) that if $\alpha_L(\nu) \rightarrow 0$, $\nu \rightarrow \infty$ then $F[\omega_\nu] \rightarrow 0$, $\omega_\nu \rightarrow -\infty$.

More generally, setting $v = \rho_*$, corresponding to a value of $\omega_v = \omega_m - \rho_* |\dot{\theta}_L|$ within the Fourier bandwidth, one can form

$$\left. \begin{aligned} & \int_{v^\dagger}^{\infty} \frac{\dot{\theta}_L F[\omega_v]}{\sqrt{v^2 - \rho_*^2}} dv = \\ & -2\pi \int_{v^\dagger}^{\infty} \frac{K_v dv}{\sqrt{v^2 - \rho_*^2}} \left(\int_{\rho^\dagger}^{\infty} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho \right) \\ & = -2\pi \int_{\rho_*}^{\infty} \frac{d \log n}{d\rho} \left(\int_{\rho_*}^{\rho} \frac{(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2)}{\sqrt{v^2 - \rho_*^2}} K_v dv \right) d\rho \\ & = \pi \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho, \quad v^\dagger = \rho_* + \hat{y}^\dagger K_{\rho_*}, \rho^\dagger = v - \hat{y}^\dagger K_v \end{aligned} \right\} \quad (6.4-13)$$

The second integral on the third line involving the Airy functions is completely deterministic. The fourth line is obtained from integrating by parts. The weighting function $W(\rho, \rho_*)$ is given by

$$\left\{ \begin{aligned} & W(\rho, \rho_*) = \\ & 2K_\rho \frac{\text{Ai}[0]^2 + \text{Bi}[0]^2}{\sqrt{\rho^2 - \rho_*^2}} - 4 \int_{\rho_*}^{\rho} \frac{\text{Ai}[\hat{y}] \text{Ai}'[\hat{y}] + \text{Bi}[\hat{y}] \text{Bi}'[\hat{y}]}{\sqrt{v^2 - \rho_*^2}} K_v^2 dv, \\ & \int_{\rho_*}^{\rho} W(\rho, \rho_*) d\rho = 2 \int_{\rho_*}^{\rho} \frac{\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2}{\sqrt{v^2 - \rho_*^2}} K_v dv \xrightarrow{\rho \rightarrow \infty} 1, \\ & \hat{y} \doteq \frac{1}{4K_v^4} (v^2 - \rho^2) \end{aligned} \right\} \quad (6.4-14)$$

In the limit as $\lambda \rightarrow 0$, $W(\rho, \rho_*)$ resembles a Dirac delta function, having similar properties. $W(\rho, \rho_*)$ and its integral, which rapidly approaches unity with increasing $\rho > \rho_*$, are shown in Fig. 6-10. The decay profile of $W(\rho, \rho_*)$ is determined by the span in \hat{y} -space over which the Airy functions make their transition to negative argument asymptotic forms, $\sim 2k^{-1}K_{\rho_*} = 2(\lambda^2 n_* r_* / 4\pi^2)^{1/3} \approx 30$ m. Therefore, $W(\rho, \rho_*)$ strongly weights the contribution from $\log n(\rho)$ in the convolution integral in Eq. (6.4-13) at $\rho = \rho_*$, and it attenuates rapidly for $\rho > \rho_*$ to nearly zero within 30 m. The half-area point is at 7 m. This spatial interval, $\sim 2(\lambda^2 n_* r_* / 4\pi^2)^{1/3}$, is where the

differences between wave theory and ray theory mainly arise. If we approximate $W(\rho, \rho_*)$ by a Dirac delta function we have

$$\frac{1}{\pi} \int_{v^\dagger(\rho_*)}^{\infty} \frac{\dot{\theta}_L F[\omega_v]}{\sqrt{v^2 - \rho_*^2}} dv = \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho \rightarrow \log n(\rho_*) \quad (6.4-15)$$

Eq.(6.4-15) is the wave theory equivalent of the Abel integral transform. The Abel transform owes its existence to a remarkable property of the integral

$$\int_a^b \frac{2x dx}{\sqrt{(x^2 - a^2)(b^2 - x^2)}} = \pi \quad (6.4-16)$$

for all real values of a and $b > a$. Unfortunately, there does not seem to be the crisp equivalent of the Abel transform in wave theory. From Eq. (6.4-13) and upon setting the Airy functions to their negative argument asymptotic forms, $\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2 \rightarrow \pi^{-1}(-\hat{y})^{-1/2}$, we have

$$2\pi \int_{\rho_*}^{\rho} \frac{(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2)}{\sqrt{v^2 - \rho_*^2}} K_v dv \rightarrow \int_{\rho_*}^{\rho} \frac{2v dv}{\sqrt{(v^2 - \rho_*^2)(\rho^2 - v^2)}} = \pi \quad (6.4-17)$$

By solving the convolution integral in Eq. (6.4-15) the refractivity profile is recovered from an integral operation on the spectral quantity $F[\omega_v]$, which is related directly to the spectral derivative of the Fourier transform of the stopped observations $\hat{E}(\rho_L, \theta_L(t))$ through Eqs. (6.3-3) and (6.4-9), and which is essentially proportional to the bending angle associated with an impact parameter value v .

The time span T used in the Fourier treatment just described has not been specified. When the SNR permits, one can partition the entire data set into a time-ordered series of contiguous subsets or data packets of temporal width ΔT_k . Over each packet a Fourier transform can be applied, and the spectrum for each of these strips can be assembled contiguously and displayed as a function of time, or as a function of nominal ray path tangency altitude, and so on. Figure 2-3, which is from [6], shows an example of this approach. This is one example of the so-called sliding spectrum technique [7].

6.4.1 Super-Refractivity

The integral for $2dG^\dagger(v)/dv$ in terms of the refractivity gradient in Eq. (6.4-6) requires special treatment to handle super-refractivity conditions. We have used $\rho = knr$ as the integration variable for convenience, but implicit in its use is the assumption that $d\rho/dr$ is positive throughout. Within a

super-refracting medium the ray curvature index, $\beta = -n'r/n$, is greater than unity. For a super-refracting spherical layer the radius of curvature of the ray is less than the radius of curvature of the refracting surface. Super-refractivity occurs when $d\rho/dr = nk(1-\beta) \leq 0$. The region where $d\rho/dr < 0$ should be evident from the data, and the critical value of dn/dr at the boundary is known, $dn/dr = -n/r \approx -157 \times 10^{-6} \text{ km}^{-1}$. Figure 6-11 shows a profile for a hypothetical refractivity gradient in the vicinity of a super-refracting layer. It could correspond to a narrow marine layer in the lower troposphere. Over the range $r_d \leq r \leq r_u$ the profile is super-refracting.

Figure 6-12 is a schematic showing the ray geometry for a super-refracting spherical layer with an upper boundary at $r = r_u$ and a lower boundary at $r = r_d$. Such layers are called ducts in ground-based radio transmission, and the word “ducting” is often used instead of “super-refracting”. Two critical rays are shown in this figure. The upper critical ray just grazes the top of the super-refracting layer; it has an impact parameter value of $n(r_u)r_u$. The lower critical ray has a tangency point at $r_* = r^c$ and an impact parameter value of $n(r^c)r^c$; it manages to escape from the top of the layer just before its ray path would have been turned inward by the strong refractive gradient in the layer. This lower critical ray escapes tangentially to the surface at $r = r_u$. It follows from Bouguer’s law that the impact parameter for this critical ray is $n(r^c)r^c = n(r_u)r_u$. Thus, in geometric optics we have a discontinuity in the bending angle profile versus impact parameter at this critical impact parameter value $\rho_* = kn(r_u)r_u$. Two rays, one just grazing the top of the layer with $r_* = r_u$, and the other at the lower critical tangency point $r_* = r^c$, both have the same impact parameter value but different bending angles. A ray with its turning point radius in the range $r^c < r_* < r_u$ can not escape; for a ray to exist the turning point radius r_* must either equal or exceed $r = r_u$, or it must be equal to or less than r^c .

Figure 6-13 shows an impact parameter curve $\rho_* = kn(r_*)r_*$ versus turning point radius r_* in the vicinity of a super-refracting layer. Note that $\rho_u < \rho_d$. A hypothetical ray with its tangency point in the range $r^c < r_* < r_u$ would have an impact parameter value in the range $\rho_u < \rho_* < \rho_d$. But from Fig. 6-13 we see that in traveling along such a ray in the range $r_* \leq r \leq r_u$, one would eventually come to a point on the ray in the super-refracting layer past which $\rho(r) < \rho_*$, which is not allowed in geometric optics for a spherical geometry. See Appendix A, Eq. (A-4b). A necessary condition for the existence of a real ray between specified endpoints is that $\rho \geq \rho_*$ at *all* points along the ray. It is no good for $\rho \geq \rho_*$ to hold part of the way; it must hold all the way between end points, or else the

term $(\rho^2 - \rho_*^2)^{1/2}$ in the bending angle and phase delay equations becomes imaginary at some point.

Figure 6-13 shows that the value of r^c relative to r_d depends on the difference $\rho_d - \rho_u$ and on the average slope of $d\rho/dr$ below the super-refracting layer. A first approximation is given by

$$r^c \approx r_d - (r_u - r_d) \frac{\bar{\beta}_{SR} - 1}{1 - \bar{\beta}^c} \quad (6.4-18)$$

where $\bar{\beta}^c < 1$ is the average value of β below the lower boundary of the super-refracting layer in the vicinity of $r = r^c$, and $\bar{\beta}_{SR} > 1$ is the average value of β in the super-refracting layer. If $\bar{\beta}^c$ is near unity, that is, nearly super-refracting, then the difference $r_d - r^c$ can be much larger than the thickness of the super-refracting layer itself.

For geometric optics, then, the refractivity profile in the range $r^c < r < r_u$ is terra incognita. Moreover, geometric optics already begins to fail⁶ before reaching these critical tangency point limits at $r_* = r^c$ and $r_* = r_u$.

No such restriction applies in wave theory, but a super-refracting layer does complicate matters. The point where $d\rho/dr = 0$ marks the boundary of a super-refractive layer, which necessitates breaking the integral in Eq. (6.4-6) for $2dG^\dagger(\nu)/d\nu$ into three sections, one section above the upper boundary at $\rho_u = kn(r_u)r_u$, one below the lower boundary at $\rho_d = kn(r_d)r_d$, and one through the layer where $d\rho/dr < 0$ ⁷. Alternatively, one can write the integral in a less convenient form in terms of the radial coordinate itself. The end point of the

⁶ For a geometric optics version of a super-refractive boundary caused by a 5% discontinuity in refractivity, see Figures 2-2 (b) and 2-8 (a). The predicted amplitude is exactly zero in the shadow zone. Figure 2-12 shows the scalar diffraction version. Figures 3-24 and 3-25 show the Mie scattering version.

⁷ Recall in Section 5.7, Eq. (5.7-27), that for a fixed spectral number ν we have set $\partial G[\rho, \nu]/\partial \rho = 0$ for $\rho \leq \rho^\dagger = \nu - \hat{y}^\dagger K_\rho$. This is an approximation that exploits the near-equality of $2dG[\rho^\dagger(\nu), \nu]/d\nu$ and $\alpha_L(\nu)$ in a medium with a moderate refractive gradient (see Appendix J). The error is small, but it depends on the curvature of the actual refractivity profile in the immediate vicinity of this turning point where the phase of the incoming wave is rapidly becoming stationary for decreasing $\rho < \nu$. See Figure 5-7 for the exact phase profile $\partial \vartheta_l^-/\partial \rho$ in an Airy layer compared to $\partial G[\rho, \nu]/\partial \rho$. In a super-refractivity zone where $\rho_u < \nu < \rho_d$, we need the integral for $2dG^\dagger(\nu)/d\nu = 2dG[\rho^\dagger(\nu), \nu]/d\nu$ only for those sections where $\rho \geq \rho^\dagger(\nu)$. Hence the three sections.

integral becomes a function of $\rho^\dagger(\nu)/kn(r^\dagger)$. This has three roots for r^\dagger when ν lies within the range $\rho_u < \nu < \rho_d$, one above, $r > r_u$, one below, $r < r_d$, and one within the super-refractive layer, $r_d < r < r_u$.

Let us define the index of refraction profile according to regime,

$$n(r) = \begin{cases} n_1(r), & r \leq r_d, \\ n_2(r), & r_d \leq r \leq r_u, \\ n_3(r), & r_u \leq r \end{cases} \quad (6.4-19)$$

Figure 6-14 is a sketch for $n(\rho)$ in the super-refracting zone. We have the continuity constraints $n_1(r_d) = n_2(r_d)$, $n_2(r_u) = n_3(r_u)$. If we now apply the wave theory version of the Abel transform in Eq. (6.4-15) to the Fourier observation function $F[\omega_\nu]$, we obtain

$$\frac{1}{\pi} \int_{\nu^\dagger(\rho_*)}^{\infty} \frac{\dot{\theta}_L F[\omega_\nu]}{\sqrt{\nu^2 - \rho_*^2}} d\nu = \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho = \quad (6.4-20)$$

$$\left. \begin{aligned} & \int_{\rho_*}^{\infty} \log n_3 W d\rho, \quad \rho_* \geq \rho_d, \\ & \int_{\rho_*}^{\rho_d} (\log n_1 - \log n_2 + \log n_3) W d\rho + \int_{\rho_d}^{\infty} \log n_3 W d\rho, \quad \rho_u \leq \rho_* \leq \rho_d, \\ & \int_{\rho_*}^{\rho_u} \log n_1 W d\rho + \int_{\rho_u}^{\rho_d} (\log n_1 - \log n_2 + \log n_3) W d\rho + \int_{\rho_d}^{\infty} \log n_3 W d\rho, \quad \rho_* \leq \rho_u \end{aligned} \right\}$$

If we replace $W(\rho, \rho_*)$ by the Dirac delta function, Eq. (6.4-20) becomes

$$\left. \begin{aligned} & \frac{1}{\pi} \int_{\nu^\dagger(\rho_*)}^{\infty} \frac{\dot{\theta}_L F[\omega_\nu]}{\sqrt{\nu^2 - \rho_*^2}} d\nu \doteq \\ & \left\{ \begin{aligned} & \log n_3(\rho_*), \quad \rho_* \geq \rho_d, \\ & \log n_1(\rho_*) - \log n_2(\rho_*) + \log n_3(\rho_*), \quad \rho_u \leq \rho_* \leq \rho_d, \\ & \log n_1(\rho_*), \quad \rho_* \leq \rho_u \end{aligned} \right\} \end{aligned} \right\} \quad (6.4-21)$$

For spectral numbers lying in the region $\rho_u \leq \nu \leq \rho_d$, only the combination $N_1(\rho_*) - N_2(\rho_*) + N_3(\rho_*)$ is recoverable with this approach.

6.4.2 Improving the Accuracy of $G^\dagger(\nu)$

One can refine the approximation in Eq. (5.7-27) for $G^\dagger(\nu)$ by forcing $G[\rho, \nu]$ to align with the exact form for the phase delay in an Airy layer. Per the discussion concerning Eqs. (5.7-24) and (5.7-25), $\vartheta_l^-(\rho)$ gives the phase delay of the l th spectral coefficient $a_l^-(\rho)$ for an incoming wave in an Airy layer. One aligns $G[\rho, \nu]$ with $\vartheta_l^-(\rho)$ at a radial distance ρ where $G[\rho, \nu]$ is still accurate. The form for $\vartheta_l^-(\rho)$ is given by

$$\vartheta_l^-(\rho) = \tan^{-1}\left(\mp \frac{\text{Bi}[\tilde{y}]}{\text{Ai}[\tilde{y}]}\right) - \tan^{-1}\left(\frac{\text{Bi}[\hat{y}]}{\text{Ai}[\hat{y}]}\right) + \text{constant} \quad (6.4-22)$$

where \tilde{y} is defined in Eq. (5.7-18) for an Airy layer, and $\tilde{y}/\hat{y} = |1 - \beta|^{-2/3}$. The top sign in Eq. (6.4-22) applies to a super-refracting medium where $\beta > 1$; the bottom sign applies when $\beta < 1$. One can readily show that the asymptotic forms for $\vartheta_l^-(\rho)$ and $G[\rho, \nu]$ for negative values of \hat{y} are identical in an Airy layer. See Fig. 5-7. For decreasing $\rho < \nu$, note that $\vartheta_l^-(\rho) - \vartheta_l^-(0)$ rapidly approaches zero.

From Eq. (6.4-22) we have

$$\frac{\partial \vartheta_l^-(\rho)}{\partial \nu} = \frac{1}{\pi K_\nu} \left(\frac{\mp \tilde{y}}{\text{Ai}[\tilde{y}]^2 + \text{Bi}[\tilde{y}]^2} - \frac{\hat{y}}{\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2} \right) \frac{1}{\hat{y}} + \text{constant} \quad (6.4-23)$$

To align $\partial G[\rho, \nu] / \partial \nu$ with $\partial \vartheta_l^- / \partial \nu$ in the vicinity of $\rho = \nu$ one sets $\partial G[\rho, \nu] / \partial \nu$ according to the following schedule

$$\frac{\partial G[\rho, \nu]}{\partial \nu} \Rightarrow \left\{ \begin{array}{l} \frac{\partial G[\rho, \nu]}{\partial \nu}, \quad \rho \geq \rho^\diamond = \nu + \hat{y}^\diamond K_\nu \\ \frac{\partial G[\rho^\diamond, \nu]}{\partial \nu} + \left(\frac{\partial \vartheta_l^-(\rho)}{\partial \nu} - \frac{\partial \vartheta_l^-(\rho^\diamond)}{\partial \nu} \right), \quad \rho \leq \rho^\diamond \end{array} \right\} \quad (6.4-24)$$

The chosen value of \hat{y}^\diamond is a compromise, taking into account on one hand the impending failure of $g(\hat{y})$ to provide the correct phase delay for $a_l^-(\rho)$ for increasing \hat{y} near zero, and on the other hand the decreasing accuracy of the Airy layer approximation to the actual refractivity profile if applied over too wide an altitude range. Eq. (6.4-23) shows that $\partial(\vartheta_l^-(\rho) - \vartheta_l^-(0)) / \partial \nu$ also rapidly approaches zero for decreasing $\rho < \nu$. Thus, we are concerned about the Airy layer approximation over a relatively narrow altitude range. The Airy layer approximation should be valid over the altitude range

$\sim -2K_v \leq \rho - \nu \leq -\hat{y}^\diamond K_v$, about 50 m. Here we have set $\hat{y}^\diamond = -2$. With this modified form the derivative of the spectral density function becomes

$$\left. \begin{aligned} \frac{dG^\dagger}{d\nu} &\Rightarrow \frac{\partial G[0, \nu]}{\partial \nu} \Rightarrow \frac{\partial G[\rho^\diamond, \nu]}{\partial \nu} + \left(\frac{\partial \vartheta_l^-(0)}{\partial \nu} - \frac{\partial \vartheta_l^-(\rho^\diamond)}{\partial \nu} \right) = \\ &\frac{\partial G[\rho^\diamond, \nu]}{\partial \nu} - \frac{1}{\pi K_v} \left(\frac{\mp \tilde{y}^\diamond}{\text{Ai}[\tilde{y}^\diamond]^2 + \text{Bi}[\tilde{y}^\diamond]^2} - \frac{\hat{y}^\diamond}{\text{Ai}[\hat{y}^\diamond]^2 + \text{Bi}[\hat{y}^\diamond]^2} \right) \frac{1}{\hat{y}^\diamond} \end{aligned} \right\} \quad (6.4-25)$$

The weighting function $W(\rho, \rho_*)$ also is slightly modified. In this regard, note that the corrective term from the Airy layer in Eq. (6.4-25) approaches zero as β approaches zero with increasing impact parameter.

The difference in the two forms for $dG^\dagger / d\nu$ in Eqs. (6.4-2) and (6.4-25) is given by

$$\left. \begin{aligned} \Delta \left(\frac{dG^\dagger}{d\nu} \right) &= -\pi K_v \int_{\rho^\dagger}^{\rho^\diamond} \frac{d \log n}{d\rho} (\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2) d\rho + \frac{\partial \vartheta_l^-(\rho^\diamond)}{\partial \nu} - \frac{\partial \vartheta_l^-(0)}{\partial \nu} \\ &\doteq \frac{d \log n}{d\rho} g(\hat{y}^\diamond) + \frac{1}{\pi \hat{y}^\diamond K_v} \left(\frac{\mp \tilde{y}^\diamond}{\text{Ai}[\tilde{y}^\diamond]^2 + \text{Bi}[\tilde{y}^\diamond]^2} - \frac{\hat{y}^\diamond}{\text{Ai}[\hat{y}^\diamond]^2 + \text{Bi}[\hat{y}^\diamond]^2} \right) \\ &\rightarrow \frac{1}{K_v} \left(\frac{\beta}{\beta-1} \right) \left(\frac{6-5\beta}{32} \right) (-\hat{y}^\diamond)^{-5/2} + O[(-\hat{y}^\diamond)^{-4}] \end{aligned} \right\} \quad (6.4-26)$$

To the extent that the Airy layer approximation is valid, this form gives the error in Eq. (6.4-2) for $dG^\dagger / d\nu$. For $\beta = 0.5$, this difference is about 0.1 mrad at $\hat{y}^\diamond = -2$; for $\beta = 2$, it is about 0.2 mrad. The error increases as $\beta \rightarrow 1$. The Airy layer analysis in Section 5.7 fails at $\beta = 1$, exactly.

With this modification one still ends up with an integration interval over ρ that is effectively truncated from below at $\rho = \nu - \hat{y}^\diamond K_v$ instead of at $\rho = \nu - \hat{y}^\dagger K_v$. A super-refractivity zone still yields three separate integration sections for $dG^\dagger(\nu) / d\nu$ when the spectral number lies in the super-refractivity zone $kn(r_u)r_u < \nu < kn(r_d)r_d$; one section above $r = r_u$, one below $r = r_u$ down to $r = r_d$, and one below $r = r_d$ down to the turning point. See also the discussion in Section 5.8 on comparison of the wave equation solutions in a super-refracting medium.

6.4.3 Resolution Issues

One could infer from Eq. (6.4-15) that away from a super-refracting region the resolution with which $n(\rho)$ may be recovered is limited only by the SNR measurement errors present in the Fourier transform quantity $F[\omega]$. But there are a number of implicit assumptions embedded in Eq. (6.4-15), not the least of which are the assumptions of spherical symmetry and error-free ionosphere effects. These and other resolution-limiting factors have been discussed in [8, 9]. Departures from spherical symmetry arise in two forms: a first part, largely deterministic, is due to departures of the geopotential of the Earth from spherical symmetry, the oblateness being the principal term. The latitude-dependent Earth flattening term is factored into the spectral formulation presented here by a small adjustment to the value of r_o and to the satellite coordinates. The second part is due to imperfect *a priori* knowledge of the topography of any given surface of constant refractivity. Along-track water vapor variability, for example. Another contributor is the geostrophic effect from winds aloft on a surface of constant pressure. *Ad hoc* calibrations could be used to correct for these usually small effects for each occultation, using some local model from ECMWF, for example; not an impossible task, but surely a tedious one. In a thin phase screen model these adjustments are equivalent to adjusting D , the limb distance, by an amount δD . It can be shown [8, 9] that an uncertainty or error δD in the adopted value of D degrades the resolution Δh , in the screen and that $\delta(\Delta h)/\mathcal{F}_o$ is at least as large as $0.45(\delta D/D)^{1/2}$, where \mathcal{F}_o is the vertical diameter of the free space first Fresnel zone, about 1.5 km. Thus, a 1% error in D maps into a limiting vertical resolution that is about 5% of \mathcal{F}_o ; a 4% error maps into about 10%, and so on. With respect to the local Fresnel limit these percentages would be greater by the factor $\xi^{-1/2}$ because of defocusing. In other words, the limiting resolution is quite sensitive to this type of error.

Similar conclusions follow for other resolution-limiting factors. The horizontal resolution is limited by the vertical resolution (see Eq. (2.2-10)). But the vertical resolution is limited in effect by the horizontal resolution, as just discussed. Equating these two expressions containing δD yields

$$\delta D \approx 90 \left(\frac{\xi}{(1-\beta)^2} \right)^{\frac{1}{3}} \text{ km} \quad (6.4-27)$$

This gives a minimum horizontal resolution of about 50 km where the defocusing is 1/10 and $\beta = 0.2$, and about 100 km where the defocusing is 1. This limiting horizontal resolution yields a limiting vertical resolution that is 15–20% of the local Fresnel diameter. That the Fresnel limit on vertical

resolution can be substantially surpassed using scalar diffraction theory techniques (30–100 m) has already been demonstrated [8, 10–12], but a realistic limit is yet to be established.

Another consideration is the cross-track and along-track drifts of the tangency point during its descent or ascent. The velocity vector of the tangency point of an occultation is rarely exactly vertical in the upper atmosphere; it is canted off-vertical depending on the orbit geometry of the satellites. The cross-track velocity of the tangency point is proportional to the tangent of this off-vertical angle. Moreover, the vertical component of the tangency point velocity vector slows because of defocusing. But, the cross-track component of the drift is not appreciably slowed and the along-track component accelerates with depth, approaching the rate $-\rho_* d\tilde{\theta}_L / dt$ in strong defocusing. These factors result in a progressively more shallow descent ratio with depth for the tangency point. In other words, the horizontal displacement of the tangency point during an occultation is a significant factor in resolution questions. Of course, in a 4DVAR context these resolution issues should be addressed in terms of the information content already extant in the 4DVAR system.

6.5 Summary

This section outlines one spectral technique based on wave theory for recovering the refractivity profile $N(\rho)$ from the spectral derivative of the Fourier transform of the received amplitude and stopped phase measurements made by the LEO over time. Coincidentally the bending angle profile $\alpha_L(\nu)$ is recovered. Spectral techniques in general facilitate recovery of bending angles and refractivity fundamentally because of their ability to uniquely sort received rays according to their excess Doppler, or impact parameter values in a spherical symmetric medium. Spectral techniques seem essential when adverse signal conditions prevail because of the concurrent reception of multiple rays. Spectral techniques also are efficacious, but perhaps not essential if a third order ray theory is used, in near-caustic situations where the validity of second order geometric optics breaks down.

Because of the close correspondence noted in Table 5-1 between the phase delay spectral density function $2G^\dagger(\nu)$ evaluated at its stationary value in spectral number and the scalar diffraction/thin screen phase function $\varphi(\rho_*)$ discussed in Chapter 2, one need not start from wave theory as the framework for obtaining these spectral results. We noted in Section 5.10 that the phase delay spectral density function $G^\dagger(\nu)$ from wave theory and the thin screen phase function $\varphi(h)$ are related by

$$-2G^\dagger(\nu) \doteq 2 \int_{\nu}^{\infty} \tilde{\alpha}(\nu, \nu) d\nu = \int_{\nu}^{\infty} \alpha_L(\nu) d\nu \Leftrightarrow \varphi(\nu) \quad (6.5-1)$$

The analog of Ψ_{Lo} in the thin screen is $\Phi(h, h_{LG})$, the Fresnel phase function given in Eq. (2.5-1). It follows that essentially the same form given in Eq. (6.4-5) would come from the scalar diffraction integral in a thin phase screen model, provide that the screen is so constructed that there is a one-to-one relationship between h and the impact parameter. Therefore, the wave-theoretic Fourier approach followed here in Chapter 6 to recover the refractivity profile can be converted into a thin screen Fourier approach by replacing $-2G^\dagger(v)$ by $\varphi(\rho_*)$ and its derivative $2dG^\dagger(v)/dv$ by $-d\varphi(\rho_*)/d\rho_*$, that is, by $\alpha_L(\rho_*)$. One also can start from a thin phase screen model, whose surface is defined by the impact parameter, and use scalar diffraction theory to obtain essentially the same results. This offers an easier approach.

Finally, an intriguing prospect for future in-receiver signal processing operations in support of limb sounding, is the incorporation of advanced processing techniques, such as essentially real-time Fourier transform algorithms. With current POD information a smart receiver not only could power through adverse signal episodes and perform backward reconstruction, it also could report bending angles for multiple rays and refractivity profiles directly, along with essential statistics, basic phase and amplitude data, and other housekeeping information.

One should not underestimate the potential cost savings of an in-receiver automated system using the GPS, especially in adverse signal conditions. For an analogy one need look no further than the navigation of a typical LEO with an on-board GPS receiver. Here one can find highly automated processes resulting in centimeter-level orbit accuracy. Once operational, the workforce required to routinely maintain and use this capability is an order of magnitude smaller than that required for most ground-based tracking systems [13]. Investment now in the necessary R&D to enable future automated in-receiver operations for GPS-based limb-sounding would seem to be a wise programmatic option. In the future other GNSS programs besides the GPS will become operational. Use of these systems also should be incorporated in future flight receiver designs.

Acknowledgement

Most of this monograph was written during 2000–2002 while I was on assignment from JPL at GeoForschungsZentrum-Potsdam (GFZ), now a leading Earth Science institution in Germany, and in the late 19th century a principal European center in the nascent field of mathematical geodesy. I wish to thank my colleagues at GFZ-Potsdam and DRL-Neustrelitz, Georg Beyerle, Jens Wickert and Norbert Jakowski, for many stimulating discussions, and to Professor Christoph Reigber, for creating an extraordinary scientific *Atmosphäre* at GFZ for visitors and staff alike. I also wish to thank the JPL

senior management for granting me this “sabbatical”. I especially wish to acknowledge support by the Solid Earth program in NASA Code Y under Dr. John LaBrecque, who played a key programmatic role in the success of GPS applications in space. To Cami Melbourne, my friend and wife, I owe a great debt for her forbearance and technical support with the preparation of text and figures. Some of the work here sprang from concepts originating in the JPL groups involved in research and development of high accuracy GPS receiver technology for flight applications, and in the radio occultation groups that use this technology. In this regard, particular thanks to George Hajj, Tom Meehan, Larry Romans, and Tom Yunck for many discussions over the years. This work was performed under NASA Contract...

References

- [1] R. Anthes, C. Rocken, and Y. H. Kuo, “Applications of COSMIC to Meteorology and Climate,” *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 115–156, 2000
- [2] T. Yunck, C. H. Liu, and R. Ware: “A history of GPS sounding,” *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 1–20, 2000
- [3] J. Thomas, “Signal-processing theory for the TurboRogue receiver,” *JPL Pub.* 95-6, Jet Propulsion Laboratory, Pasadena, CA, 1995
- [4] T. Meehan (on BJ operations xx)
- [5] J. Jackson, *Classical Electrodynamics*, 2nd Ed., J. Wiley & Sons, Inc., 1975
- [6] G. Lindal, J. Lyons, D. Sweetnam, V. Eshelman, D. Hinson and G. Tyler, “The atmosphere of Uranus: Results of radio occultation measurements with Voyager 2”, *J. GeoPhy. Res.*, vol. 92, pp. 14987–15001, 1987
- [7] S. Sokolovskiy, “Modeling and inverting radio occultation signals in the moist troposphere,” *Rad. Sci.*, vol. 36 (3), pp. 441–458, 2001
- [8] E. Marouf, G. Tyler and P. Rosen, “Profiling Saturn’s rings by radio occultation,” *Icarus*, vol. 68, pp. 120–166, 1986
- [9] W. Melbourne, E. Davis, C. Duncan, G. Hajj, K. Hardy, E. Kursinski, T. Meehan, L. Young and T. Yunck, “The application of spaceborne GPS to atmospheric limb sounding and global change monitoring,” *JPL Publication 94-18*, Jet Propulsion Laboratory, Pasadena, CA, 1994
- [10] E. Karayel and D. Hinson, “Sub-Fresnel scale vertical resolution in atmospheric profiles from radio occultation,” *Radio Sci.*, vol. 32 (2), pp. 411–423, 1997
- [11] M. Gorbunov, A. Gurvich, and L. Bengtsson, “Advance algorithms of inversion of GPS/MET satellite data and their application to

- reconstruction of temperature and humidity,” *Max Planck Inst. für Meteor.*, Rpt. No. 211, Hamburg, Germany, 1996
- [12] K. Igarashi, K. Pavelyev, J. Wickert, K. Hocke, and D. Pavelyev, “Application of radio holographic method for observation of altitude variations of the electron density in the mesosphere/lower thermosphere using GPS/MET radio occultation data,” *J. Atmos. Sol.-Terr. Phys.*, vol. 64, pp. 959-969, 2002
- [13] Melbourne et al 1996? (Hartl 65th birthday xx)

Figure 6-1. Hypothetical bending angle at the LEO for a layer of higher refractivity embedded in dry air near sea level. Figure 6-7 shows the recovered refractivity profile for this water vapor layer.

Figure 6-2. Impact parameter diagram for the refractive bending angle profile shown in Figure 6-1.

Figure 6-3. Amplitude of the field at the LEO from the bending angle profile given in Figure 6-1.

Figure 6-4. Phase of the complete field at the LEO minus the phase of the main ray (m) resulting from the Figure 6-1 bending angle profile.

Figure 6-5. Blow-up of Figures 6-3 and 6-4 at $\theta_L - \theta_{\rho_o} \approx 2.1$, showing the burst of phase acceleration at $\theta_L - \theta_{\rho_o} = 2.12$ resulting from near-complete cancellation of the field by opposing caustic rays.

Figure 6-6. Multipath with single ray paradigm.

Figure 6-7. Error in recovered refractivity $\hat{N}(\rho_)$ as a result of ignoring the dashed sections of the (m) and (a) rays in Figure 6-6.*

Figure 6-8. Spherical geometry for non-coplanar LEO and GPS orbits. GPS satellite located infinitely afar along the negative z -axis.

Figure 6-9. Obliquity factor $d\tilde{\theta}_L / d\theta_L$ and I / ε versus LEO orbit plane inclination angle I .

Figure 6-10. Weighting function $W(\rho, \rho_)$ in the wave theory analog of the Abel transform.*

Figure 6-11. Gradient of $n(r)$ near a super-refracting medium.

Figure 6-12. Ray geometry for a super-refracting spherical layer.

Figure 6-13(L). Impact parameter curve in a super-refracting zone.

Figure 6-14(R). Profile for n versus ρ in a super-refracting zone.

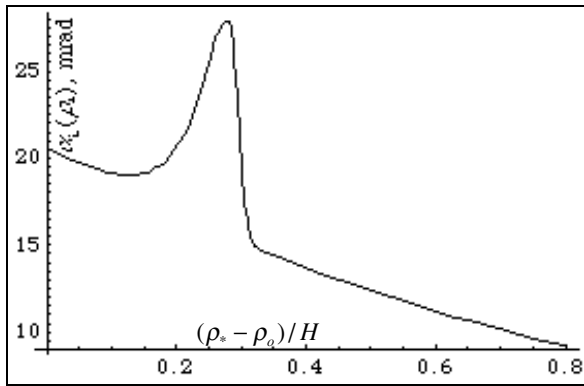


Figure 6-1.

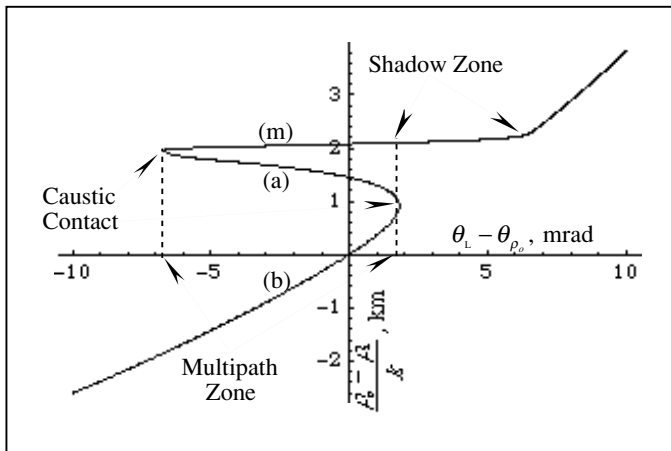


Figure 6-2.

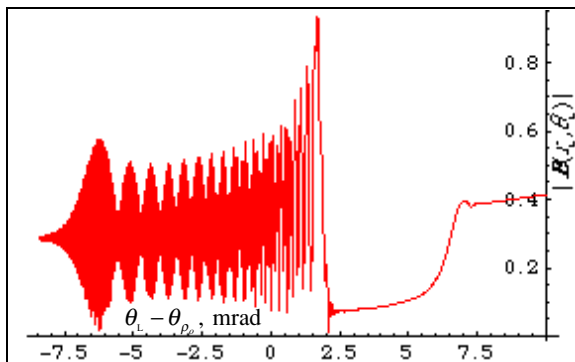


Figure 6-3.

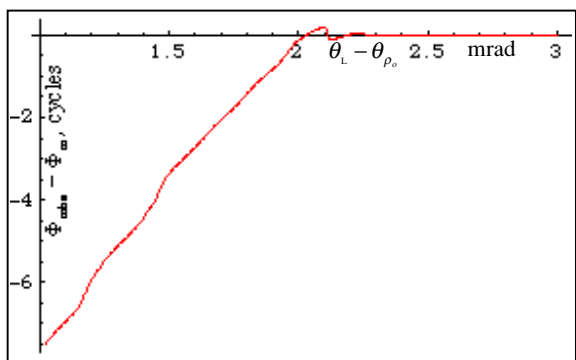


Figure 6-4.

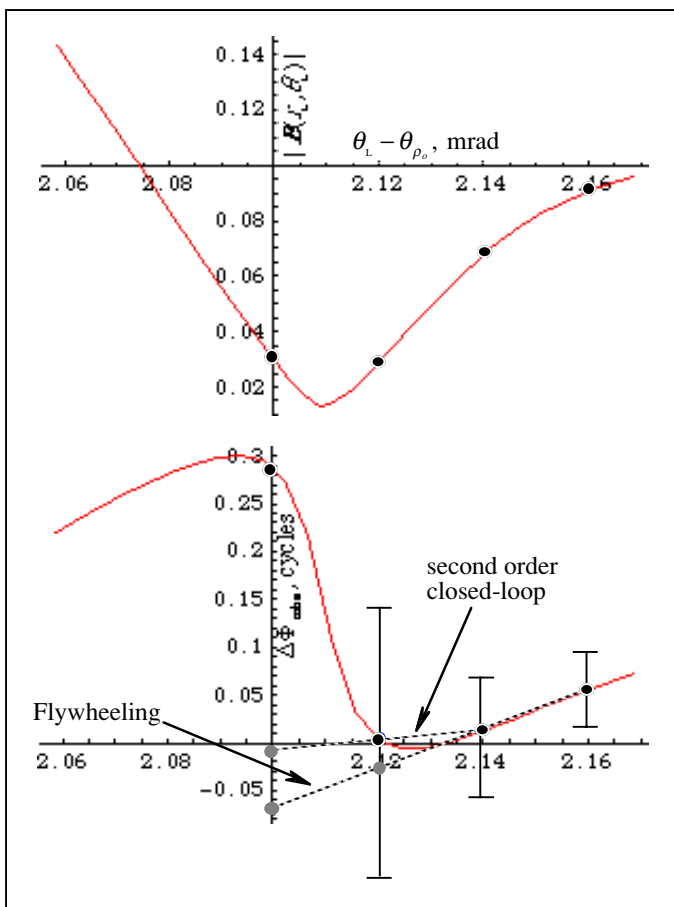
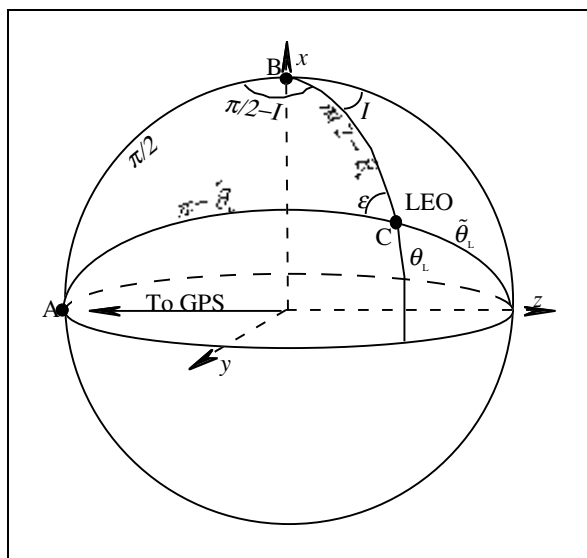
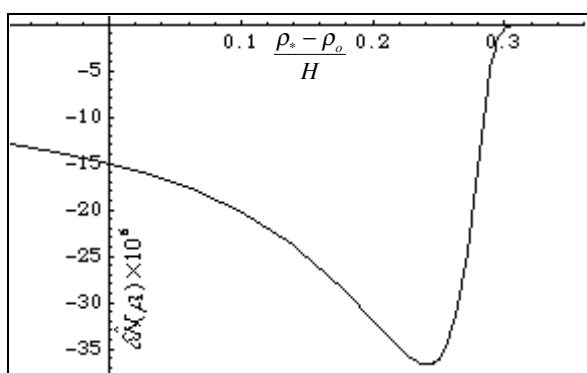
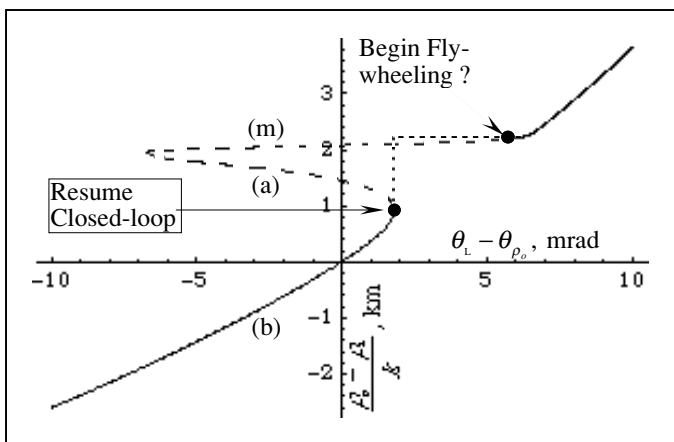


Figure 6-5.



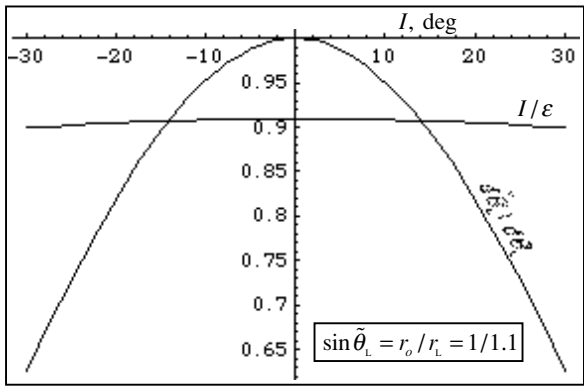


Figure 6-9.

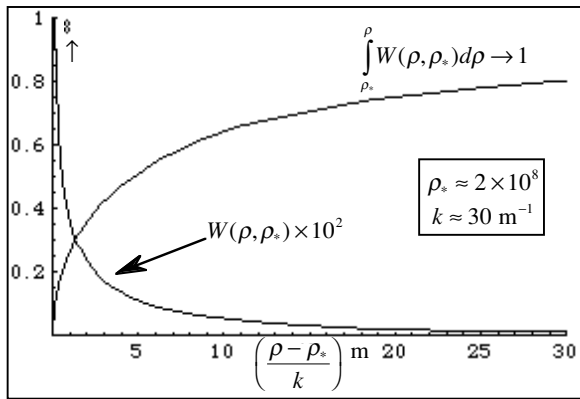


Figure 6-10.

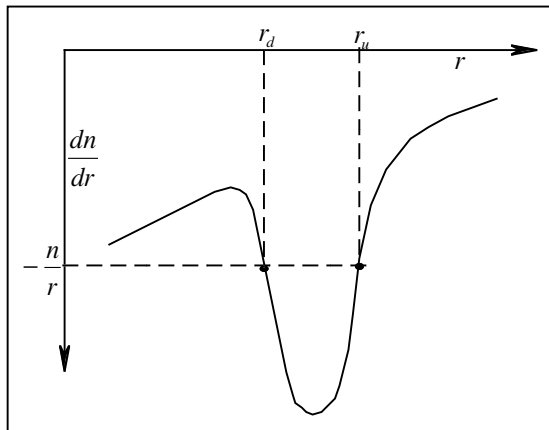


Figure 6-11.

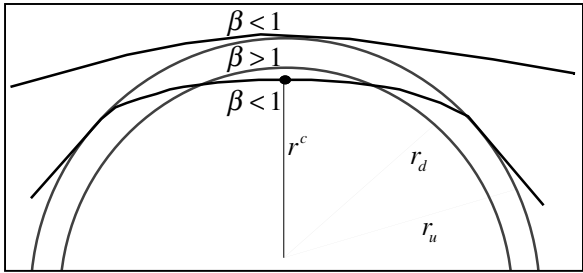


Figure 6-12.

Figure 6-13(L).

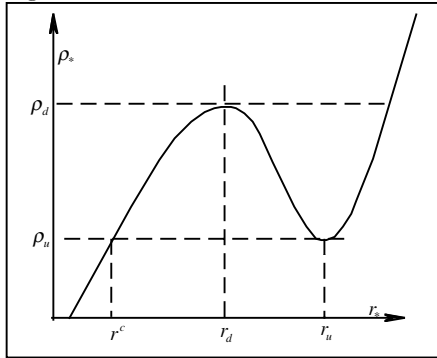


Figure 6-14(R).

